



Review

Current Applications of Artificial Intelligence in the Neonatal Intensive Care Unit

Dimitrios Rallis ¹, Maria Baltogianni ¹, Konstantina Kapetaniou ² and Vasileios Giapros ^{1,*}

¹ Neonatal Intensive Care Unit, School of Medicine, University of Ioannina, 45110 Ioannina, Greece; drallis@uoi.gr (D.R.); mbalt@doctors.org.uk (M.B.)

² Department of Pediatrics, School of Medicine, University of Ioannina, 45110 Ioannina, Greece; k.kapetaniou@uoi.gr

* Correspondence: vgiapros@uoi.gr; Tel.: +30-2651-099-326

Abstract: Artificial intelligence (AI) refers to computer algorithms that replicate the cognitive function of humans. Machine learning is widely applicable using structured and unstructured data, while deep learning is derived from the neural networks of the human brain that process and interpret information. During the last decades, AI has been introduced in several aspects of healthcare. In this review, we aim to present the current application of AI in the neonatal intensive care unit. AI-based models have been applied to neurocritical care, including automated seizure detection algorithms and electroencephalogram-based hypoxic-ischemic encephalopathy severity grading systems. Moreover, AI models evaluating magnetic resonance imaging contributed to the progress of the evaluation of the neonatal developing brain and the understanding of how prenatal events affect both structural and functional network topologies. Furthermore, AI algorithms have been applied to predict the development of bronchopulmonary dysplasia and assess the extubation readiness of preterm neonates. Automated models have been also used for the detection of retinopathy of prematurity and the need for treatment. Among others, AI algorithms have been utilized for the detection of sepsis, the need for patent ductus arteriosus treatment, the evaluation of jaundice, and the detection of gastrointestinal morbidities. Finally, AI prediction models have been constructed for the evaluation of the neurodevelopmental outcome and the overall mortality of neonates. Although the application of AI in neonatology is encouraging, further research in AI models is warranted in the future including retraining clinical trials, validating the outcomes, and addressing serious ethics issues.

Keywords: artificial intelligence; machine learning; deep learning; neonate



Citation: Rallis, D.; Baltogianni, M.; Kapetaniou, K.; Giapros, V. Current Applications of Artificial Intelligence in the Neonatal Intensive Care Unit. *BioMedInformatics* **2024**, *4*, 1225–1248. <https://doi.org/10.3390/biomedinformatics4020067>

Academic Editors: Alexandre G. De Brevern and James C. L. Chow

Received: 30 March 2024

Revised: 21 April 2024

Accepted: 7 May 2024

Published: 9 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

AI (artificial intelligence) refers to computer algorithms that replicate the cognitive function of humans, using specified operational models produced from the statistical assessments of big data sets [1]. During the last decades, AI has been applied in several aspects of human life, including the healthcare industry [2]. This accomplishment has been made possible by updated hardware technologies and ever-more-complex computer algorithms for processing and storing massive datasets [3–6]. Across healthcare fields, however, there seems to be varying degrees of enthusiasm for the topic of AI research. Nearly half of the evidence arises from published studies in the adult medical sciences (pathology, oncology, neurology, cardiology, gastroenterology, dermatology, pulmonology, endocrinology, emergency medicine), followed by imaging sciences (cell imaging, radiology), and by studies in surgery, ophthalmology, psychiatry, and pediatrics (Figure 1) [3,6,7].



Figure 1. Studies on artificial intelligence by medical specialty. Based on evidence from references [6,7].

Although AI has enormous potential, its use in neonatology is still in its early stages. Since the pioneer studies of AI application to neonatal neuromonitoring during the decade of 1990s [8,9], AI-based systems have been gradually expanded to the diagnosis and management of common neonatal morbidities including but not limited to respiratory distress syndrome (RDS) [10,11], bronchopulmonary dysplasia (BPD) [12–17], patent ductus arteriosus (PDA) [18,19], neonatal sepsis [20,21], and retinopathy of prematurity (ROP) [22–30]. Furthermore, AI-based algorithms are on a daily basis used for the prediction of long-term neonatal outcomes and mortality [31–45]. Despite this evolution, however, there are concerns about how AI will be incorporated into the healthcare system because there is a growing demand for early detection, alarm systems, and diagnostic testing [46]. Compared to earlier, there are now higher expectations for AI in their daily practices; besides, evidence from previous studies underscores the limitations and risks of AI applications including several ethical concerns.

Reviews of AI application to neonatal monitoring are scarce and the aim of this review is to cover this very important issue. Moreover, as several novel practices, as AI, are first applied to adult or pediatric populations, this review also aims to motivate neonatologists to seek further information and co-operation with other scientists to explore the perspectives of AI in this very crucial period of life. In this narrative review, we evaluate AI's current applications and advantages in the neonatal intensive care unit (NICU) and explore the perspectives of AI on neonatal care in the future. We, therefore, examined the existing evidence of AI-based monitoring and diagnostic tools that could support the care and follow-up of neonatologists. We explore several AI designs for image, signal, and electronic health record processing, evaluate the benefits and drawbacks of recently

developed decision support systems, and shed light on potential future applications for physicians and neonatologists in their routine diagnostic work.

Our study is organized into (1) presenting the basic AI models applied in neonatal care, (2) grouping AI applications that pertain to neonatology into domains, elucidating their sub-domains, and highlighting the key components of the relevant AI models, (3) reviewing and providing a thorough summary of the latest research with a focus on applying AI to all areas of neonatology, and (4) examining and discussing the existing challenges related to AI in neonatology, as well as directions for future study (Figure 2).

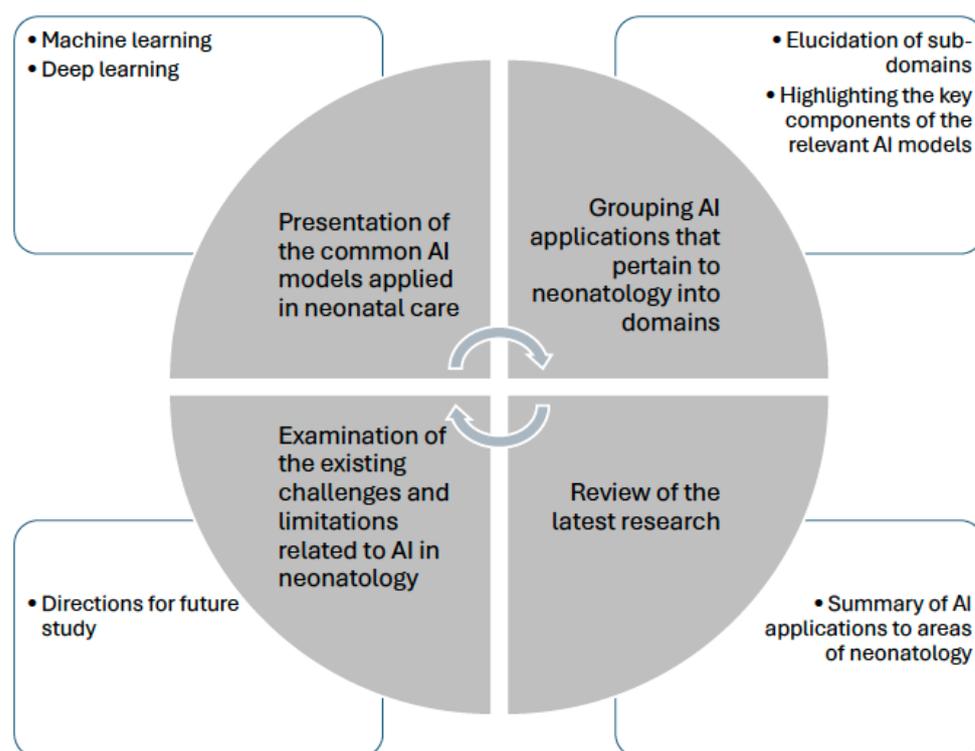


Figure 2. Overview of the study organization.

2. Basic Models of Artificial Intelligence

The AI framework is based on machine learning (ML) and deep learning (DL), two subsets of AI that have been widely applied to the healthcare industry [47]. To create models based on datasets that enable the algorithm to generate predictions and make judgments without programming, ML refers to the automatic improvement of AI algorithms through experience and vast amounts of historical data. ML uses both unstructured data that are difficult to arrange using predetermined structures (e.g., clinical notes), as well as structured data that are easily organized into predefined structures. Furthermore, ML models generate software algorithms to develop AI decision-support systems [47]. The majority of these systems are created using standard algorithms, which consistently produce the same outcome for a given input, and thus, decision-support systems help healthcare professionals analyze enormous amounts of information [48–50].

Unlike this broader definition of ML, the fundamental idea behind DL is derived from the neural networks of the human brain that process and interpret information. To simulate this process, DL is based on representation learning and artificial neural networks (ANNs), and when the number of layers is large (i.e., deep) simulates more intricate links between input and output [51,52]. The ANN is a mathematical model that mimics the composition and operation of biological neural networks. The quantity and configuration of an ANN's neural layers as well as the training set determine its performance [53]. The main subtypes of DL networks are convolutional neural networks (CNNs), recurrent neural networks (RNNs), and generative adversarial neural networks (GANs) [54]. CNNs are mostly utilized

in computer vision and signal processing applications. The CNN architecture consists of a series of stages, or layers, that make it easier to obtain hierarchical characteristics. Later phases extract more global characteristics, while initial phases extract more local features, like corners, edges, and lines. As features spread from one layer to another, the representation of those features is richer [55]. In medicine, CNNs are most commonly employed for image processing and detection, especially in radiology, pathology, and dermatology [55]. RNNs are more effective when handling time-series data, such as clinical data or electronic health records (EHRs), and sequential data, such as text and speech [56]. GANs are a subtype of the DL model that can be used to create new data that is similar to existing data [57]. Finally, natural language processing (NLP) is an AI technology that aids computers in comprehending and interpreting human language, organizing clinical notes and unstructured data, and thus, enabling better decision-making [58,59]. Figure 3 presents the basic models of AI.

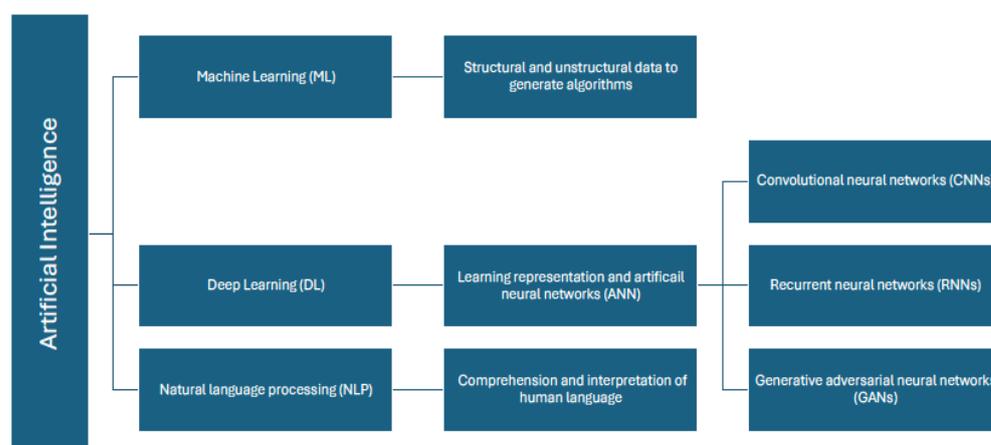


Figure 3. Basic models of artificial intelligence.

3. Domains of Artificial Intelligence's Applications in Neonatal Care

3.1. Neuromonitoring

The previous decades have seen increased research on the neuromonitoring of critically ill neonates, thanks to the advancements in AI (Table 1). AI, and especially ML, has made it possible for computer systems to examine and analyze massive amounts of data, including medical patterns, mainly applied to the electroencephalogram (EEG) and magnetic resonance imaging (MRI) [60].

3.1.1. Electroencephalography

Seizures are the most common neurological emergency in the neonatal population, and most likely occur during the first days of life [61]. Seizures are more common in neonates born at less than 30 and more than 36 weeks of gestation, with the frequency of seizures in neonates estimated to be around 8% [61]. Additionally, evidence suggests that treating seizures early on enhances the patient's response to medication [62], while it is well known that recurrent seizures are linked to worse long-term neurodevelopmental outcomes, regardless of the underlying cause [63,64]. Seizures are particularly difficult to diagnose in the neonatal population because they can be difficult to distinguish from normal infant movements even when they do occur, or they can be limited to electrographic episodes [65]. Although neonatal seizures need to be treated right away, it can be extremely challenging to recognize, since up to 85% of neonatal seizures may not have any clear clinical symptoms.

In the NICU, EEG has emerged as a crucial component of neurocritical care, as it is crucial to identify neonatal seizures and allows the distinction between epileptic seizures and nonepileptic episodes [66]. Additionally, EEG monitoring helps uncouple clinical and EEG seizures after antiseizure treatment [67], detecting the electrical discharge that

may persist after therapy, while the clinical manifestation of the seizure that may have existed before treatment disappeared. EEG records non-invasively the electrical activity of the cerebral cortex allowing for the real-time evaluation of cortical background function; however, real-time review and implementation of EEG can be challenging. Moreover, continuous EEG (cEEG) increases the diagnostic and prognostic potential, since it allows the evaluation of the background activity over time [68]. Thus, cEEG monitoring is the recommended standard of care for identifying and treating all seizures quickly [69,70]. Due to the challenges in acquiring traditional EEG, NICUs have currently adopted a less precise but more straightforward method of EEG monitoring, the amplitude-integrated EEG (aEEG). As opposed to cEEG monitoring, aEEG is a bedside device that shows one or two channels of filtered, smoothed, and quantitatively converted EEG data, while the cortical electrical activity is compressed in duration and converted in a semi-logarithmic chart [71,72]. However, aEEG does not have an ideal sensitivity, specificity, and interobserver agreement for identifying seizures [73], and thus, it is recommended to serve as an adjunct to cEEG monitoring [68,74].

During the past few decades, research in AI, and particularly DL, has evolved in the field of the creation of automatic seizure detection algorithms [75]. These algorithms exhibit remarkable seizure detection accuracy, comparable to that of human specialists [76]. In 1992, Liu et al. proposed a computerized detection system for neonatal seizures, and thereafter, numerous methods have been documented, refined, and verified [8]. The performance of the initial automatic seizure detection algorithms was suboptimal for therapeutic use as they had been created by modifying algorithms intended for adult users [8,9]; however, to date many seizure detection algorithms have been developed mainly for full-term but also preterm neonates [77]. The development of these algorithms requires the labeling of seizures by several specialists as well as obtaining enough data for testing, training, and validation.

In 2020, a randomized clinical trial assessing the effect of ML on the real-time identification of neonatal seizures in a NICU was published [78]. According to that report, more seizures were recognized in real-time, when AI algorithms were applied in the NICU [78]. Following extensive training and offline analysis, the accuracy of the recognition of electrographic seizures both with and without the automatic seizure detection algorithms was tested in a multicenter clinical trial, suggesting that the algorithm could serve as a bedside tool in clinical practice [79,80]. The model greatly enhanced the recognition of seizure hours, even though the set aim of improving the detection of specific neonates with seizures was not fulfilled.

In addition to monitoring and treating newborn seizures, EEG is also a valuable diagnostic tool for neonatal encephalopathy, namely hypoxic-ischemic encephalopathy (HIE). AI research is being conducted to create algorithms, many of which use DL techniques, that can evaluate brain maturation, estimate sleep stages [81], and grade background EEG patterns in HIE [82]. Automated EEG interpretation based on ML technology has recently shown good performance in detecting HIE severity and can be helpful in the early severity grading of neonatal HIE [83,84]. Such an example of advanced signal processing included the convolutional neural network structures, which can self-extract convolutional features from raw EEGs [82]. Besides, the possible application of AI in predictive modeling for electrographic seizures in newborns with HIE was examined by Pavel et al., with the goal of early detection of infants most at risk of recurrent seizures [85]. ML algorithms were created for clinical and both qualitative and quantitative EEG characteristics. Notably, both the automated quantitative EEG analysis and the analysis carried out by a skilled neurophysiologist (qualitative) increased the predictive value of these models by incorporating clinical data. These studies highlight the possibility of using ML in evaluating the EEG background of neonates with HIE.

3.1.2. Magnetic Resonance Imaging

The application of AI to enhance the utility and inference from brain MRI has advanced significantly during the last few years. Technical advancements in AI techniques include methods to reduce movement artifact effects and boost information yield, as well as advancements in tissue classification [86]. These have made it possible for a deeper evaluation of the developing brain, and a new understanding of the effects of prenatal events on structural and functional network topologies [87].

One of the regions in the neonatal brain where myelination starts is the posterior limb of the internal capsule (PLIC). Crucially, both term and preterm newborns' neurological outcomes depend on the proper and timely maturation of the PLIC. Abnormalities in the PLIC detected on MRI have been linked to hemiplegia, and worse neurodevelopmental outcomes [88]. Over the past few decades, there has been a noticeable rise in the prevalence of cerebral palsy to over 2.0 per 1000 live births, which is inversely proportional to the gestational age and carries significant lifetime burdens [89,90]. An ML algorithm for the automated segmentation and quantification of the PLIC in preterm newborns undergoing MRI was proposed in a recent work [91], where authors demonstrated good accuracy for the ML model when compared to expert analysis, indicating the successful application of their algorithm to a large dataset. Although promising, it is necessary to evaluate how well this approach will work in clinical settings.

Identifying neuroanatomic phenotypes and predicting the outcome are the major areas in the clinical domain where AI is facilitating innovation. Preterms are characterized by a specific phenotype including abnormal brain development, cerebral palsy, autism spectrum disorder, attention deficit hyperactivity disorder, psychiatric illness, and issues with language, behavior, and socioemotional functions [92]. Abnormalities of structural and functional networks are frequent in preterm neonates as they have been obtained from structural, diffusion, and functional MRI [93]. Models that combine data from two or more imaging modalities into a single framework, can reveal previously unknown patterns of neuroanatomic variants in preterm neonates that are related to cognitive and motor outcomes [94]. Diffusion tensor metrics, neurite orientation dispersion, regional volumes, and density imaging measurements are among the several forms of MRI data that are integrated into a single model to compute morphometric similarity networks [95]. This kind of research helps identify the neural roots of cognition and behavior, identify the networks that most contribute to atypical brain development, and examine the drivers of brain dysmaturation and resilience.

Current research also aims to compare traditional computer vision approaches with efficient networks that generate reliable and accurate segmentation. To evaluate methods for segmenting newborn tissue, T1W, and T2W pictures were provided with manually segmented structures; segmenting myelinated from unmyelinated white matter is, nevertheless, still challenging [96]. The limited number of high-quality labeled data must also be acknowledged as a key limitation when comparing earlier attempts on newborn brain segmentation [97].

Table 1. Examples of the current evidence of artificial intelligence application in neuromonitoring in neonatology.

Aim	References	Artificial Intelligence Model	Data-Set Analyzed	Outcome
Electroencephalography				
Automated seizures detection	O'Shea et al. [76]	DL detection models based on SVM system and AUROC	Continuous EEG recordings	The system achieved a 56% relative improvement, reaching an AUROC of 98.5%; this compared favourably both in terms of performance and run-time
	Liu et al. [8]	SAM analysis	Continuous EEG recordings	SAM analysis demonstrated a sensitivity of 84% and a specificity of 98% in effectively differentiating between EEG epochs containing seizures and those without

Table 1. Cont.

Aim	References	Artificial Intelligence Model	Data-Set Analyzed	Outcome
Severity grading of neonatal HIE	Gotman et al. [9]	Spectral analysis	Continuous EEG recordings (281 h of recordings containing 679 seizures)	71% of the seizures and 78% of seizure clusters were detected, with a false detection rate of 1.7/h The algorithm had an AUROC of 88.3% when tested on preterm as compared to 96.6% when tested on term EEG. When re-trained on preterm EEG, the performance increased to 89.7%. An alternative DL approach showed a more stable trend when tested on the preterm cohort, starting with an AUROC of 93.3% for the term-trained algorithm and reaching 95.0% by transfer learning from the term model using available preterm data
	O'Shea et al. [77]	DL detection models based on SVM system	Continuous EEG recordings	Sensitivity and specificity were 81.3% and 84.4% in the algorithm group compared to 89.5% and 89.1% in the non-algorithm group, respectively; the false detection rate was 36.6% in the algorithm group and 22.7% in the non-algorithm group. The percentage of seizure hours correctly identified was higher in the algorithm group than in the non-algorithm group (difference 20.8%)
	Pavel et al. [78]	Algorithm for automated neonatal seizure recognition	Continuous EEG recordings	SDA achieved seizure detection rates of 52.6–75.0%, with false detection rates of 0.04–0.36 FD/h. Time based comparison of expert and SDA annotations using Cohen's Kappa Index revealed a best performing SDA threshold of 0.4 (Kappa 0.630)
	Mathieson et al. [80]	SDA	Continuous EEG recordings	The 4 grade AGS had a classification accuracy of 83% compared to human annotation of the EEG. EEG-only measures were shown to be less effective in grading the EEG than features estimated on the created sub-signals, and performance was further enhanced by adding more sub-grades based on EEG states to the AGS
	Stevenson et al. [79]	ML classifier models of AGS based on a multi-class linear analysis and AUROC	Continuous EEG and clinical data	The proposed EEG HIE-grading system achieved an accuracy of 88.9% and kappa of 0.84 on the development dataset. Accuracy for the large unseen test dataset was 69.5% and kappa of 0.54, which is a significant ($p < 0.001$) improvement over a state-of-the-art feature-based method with an accuracy of 56.8% and kappa of 0.39
	Raurale et al. [82]	Quadratic time-frequency distribution with a CNN	EEG data	The optimal solution had a 97% classification accuracy overall, ranging from 81 to 100% across the subjects
	Moghadam et al. [83]	SVM, multilayer feedforward neural network or RNN	EEG data (13,200 5-min EEG epochs)	Effective parameterization of continuous EEG data has been achieved resulting in high classification accuracy (89%) to grade background EEG abnormalities
Matic et al. [84]	Automated algorithm to quantify background EEG abnormalities	Continuous EEG recordings of 1 h	Low Apgar, need for ventilation, high lactate, low base excess, absent sleep-wake cycle, low EEG power, and increased EEG discontinuity were associated with seizures. The following predictive models were developed: clinical (MCC 0.368, AUC 0.681), qualitative EEG (MCC 0.467, AUC 0.729), quantitative EEG (MCC 0.473, AUC 0.730), clinical and qualitative EEG (MCC 0.470, AUC 0.721), and clinical and quantitative EEG (MCC 0.513, AUC 0.746). The clinical model by itself performed much worse than the clinical and qualitative-EEG model (MCC 0.470 vs. 0.368, p -value 0.037). With a p -value of 0.012, the clinical model was significantly surpassed by the quantitative-EEG model and clinical model (MCC 0.513 vs. 0.368). Performance for quantitative aEEG was MCC 0.381, AUC 0.696 and clinical and quantitative amplitude EEG was MCC 0.384, AUC 0.720	
Pavel et al. [85]	ML models (random forest and gradient boosting algorithms) using MCC and AUROC	Clinical and EEG parameters at <12 h of birth	The model significantly outperforms state-of-the-art neonatal quiet sleep detection algorithms, with mean Kappa 0.77 ± 0.01 (with 8-channel EEG) and 0.75 ± 0.01 (with a single bipolar channel EEG)	
Sleep stage classification	Ansari et al. [81]	CNN inception block	EEG data	
Magnetic Resonance Imaging				
Automated segmentation and quantification of the PLIC	Gruber et al. [91]	CNN-based pipeline comprised of slice-selection modules and a multi-view segmentation model	MRI volume data	The proposed method was capable of identifying a specific desired slice from the MRI volume

Table 1. Cont.

Aim	References	Artificial Intelligence Model	Data-Set Analyzed	Outcome
Combination of structural and functional networks	Ball et al. [94]	A data-driven, multivariate approach that integrated several imaging modalities	Clinical factors and MRI data	Five independent patterns of neuroanatomical variation that related to clinical factors included age, prematurity, sex, intrauterine complications, and postnatal adversity. It was established that there was a connection between poor cognitive and motor outcomes at two years old and imaging indicators of neuroanatomical abnormalities
	Galdi et al. [95]	Morphometric similarity networks	MRI data, such as density imaging metrics, neurite orientation dispersion, regional volumes, and diffusion tensor metrics	The regression model predicted post-menstrual age at scan with a mean absolute error of 0.70 ± 0.56 weeks; the classification model achieved 92% accuracy
Generate reliable and accurate segmentation	Makropoulos et al. [96]	A system for precisely segmenting the developing neonatal brain based on intensity	MRI data	Across a broad range of gestational ages, from 24 weeks gestational age to term-equivalent age, the suggested approach produced extremely accurate results The best test mean DSC values that were statistically significant were obtained by the dual-modality HyperDense-Net. For all tissue types, T2-weighted image processing performed better by the single-modality LiviaNET than T1-weighted image processing. Both neural networks achieved previously reported performance
	Ding et al. [97]	DSC for each tissue type against eight test subjects	MRI data	

DL, deep learning; AUROC, area under the receiver operating characteristic curve; EEG, Electroencephalogram; SAM, Scored autocorrelation moment; SVM, support vector machine; SDA, seizure detection algorithm; HIE, hypoxic-ischemic encephalopathy; ML, machine learning; AGS, automated grading systems; CNN, convolutional neural network; RNN, recurrent neural network; MCC, Matthews correlation coefficient; PLIC, posterior limb of internal capsule; MRI, magnetic resonance imaging; DSC, dice similarity coefficient.

3.2. Neurodevelopmental Outcome

ML techniques have been widely used for the neurodevelopmental evaluation and follow-up of preterm neonates (Table 2). Numerous studies used ML techniques to examine brain connections [40,98–100], brain structure analysis, and brain segmentation in preterm neonates [45,101]. Evidence suggests an association between lower brain volume, cortical folding, axonal integrity, and microstructural connectivity with preterm birth [41,102]. Additional effects of prematurity on the developing connectome have been found in studies examining functional markers of brain maturation [40,103].

Neurocognitive assessments are among the most significant domains of neurodevelopment outcomes at two years of age. Previous studies assessed how the brain's morphological alterations relate to neurocognitive outcomes [39,43,44] and the prediction of brain age [104]. It has been demonstrated that multivariate models combining near-term structural MRI findings and white matter microstructure on diffusion tensor imaging may help identify preterm neonates at risk for language impairment and guide early intervention [43,44]. Moreover, to predict neurodevelopmental impairment at two years of age, a self-training deep neural network model has been suggested, using MRI data obtained in very preterm neonates at term-equivalent age [31]. Besides, according to a study that used ML techniques to assess the impact of *PPAR* gene activity on brain development, a significant correlation was found between aberrant brain connectivity and *PPAR* gene signaling's role in aberrant white matter development [105].

ML models have been used to evaluate the association of the developmental outcome regarding language skills with the near-term MRI findings in previous studies. By examining MRI characteristics and perinatal clinical data, Valavani et al. employed ML to predict language skills at two years of corrected age in preterm neonates [42]. Language delay could be accurately predicted by delayed myelination patterns and specific clinical characteristics. The authors concluded that ML models could be useful for healthcare services and enhance the long-term outcomes of preterm neonates. Furthermore, in a recent study, Balta et al. proposed an AI-based automated monitoring of newborns' general motions, a crucial screening test for detecting neuromotor problems in children [33]. The authors created an automated model to analyze infants' overall motions, by processing

videos taken with a simple camera at home. Certain patterns of spontaneous movements, such as the absence of fidgety movements or the presence of predominately contracted coordinated movements, were particularly indicative in predicting cerebral palsy in infants between the ages of 3 and 5 months of age [33].

Table 2. Examples of the current evidence of artificial intelligence application in neonatal neurodevelopmental outcome.

Aim	References	Artificial Intelligence Model	Data-Set Analyzed	Outcome
Detection of neonates with cognitive impairment	Wee et al. [44]	Clustering coefficients of individual structures were computed. SVM and canonical correlation analysis	Diffusion tensor imaging tractography and neurodevelopmental scales	At 24 months of age, the right amygdala’s clustering coefficient was linked to both internalising and externalising behaviours; at 48 months of age, the right inferior frontal cortex and insula’s clustering coefficients were linked to externalising behaviours SNPs with expected effects such as protein coding and nonsense-mediated decay were found predominantly in introns or regulatory regions of <i>PPARG</i> , where they were significantly overrepresented. The <i>PPARG</i> signaling has a previously unrecognized role in cerebral development. The proposed model achieved an accuracy of 71.0%, a specificity of 71.5%, a sensitivity of 70.4% and AUROC of 0.75, significantly outperforming transfer learning models through pre-training approaches. Based on regional white matter architecture on diffusion tensor imaging, infants at high risk for language impairments were predicted with good accuracy (sensitivity, specificity) for expressive (100%, 90%), receptive language (100%, 90%), and composite (89%, 86%) language. The model achieved balanced accuracy 91%, sensitivity 86%, and specificity 96%. As the values of the radial diffusivity, axial diffusivity, and peak width of skeletonized fractional anisotropy determined from diffusion MRI increased, the likelihood of language delay at two years of corrected age increased as well.
	Krishnan et al. [105]	ML using Sparse Reduced Rank Regression	Whole-brain diffusion tractography together with genomewide, SNP-based genotypes and neurodevelopmental scales	
	Ali et al. [31]	Self-training deep neural network	Brain functional connectome and cognitive assesment data	
Detection of neonates at risk of language impairment	Vassar et al. [43]	Multivariate models with leave-one-out cross-validation and exhaustive feature selection	MRI and white matter microstructure assessed on diffusion tensor imaging and neurodevelopmental scales	The model achieved balanced accuracy 91%, sensitivity 86%, and specificity 96%. As the values of the radial diffusivity, axial diffusivity, and peak width of skeletonized fractional anisotropy determined from diffusion MRI increased, the likelihood of language delay at two years of corrected age increased as well.
	Valavani et al. [42]	Feature selection and a random forests classifier	MRI data and neurodevelopmental scales	
Detection of neuromotor problems and risk of cerebral palsy	Balta et al. [33]	Tracking software of DeepLabCut using a k-means algorithm	Single commercial videos of six Poles on the infant’s upper body: left and right shoulders, elbows, and wrists	The results demonstrated that gross motor metrics may be meaningfully estimated and potentially used for early identification of movement disorders

ML, machine learning; SNP, single-nucleotide polymorphism; AUROC, area under the receiver operating characteristic curve; MRI, magnetic resonance imaging.

3.3. Respiratory System

One of the main causes of infant mortality and morbidity in preterm deliveries is BPD. Although several biomarkers have been associated with the emergence of RDS, there are currently no meaningful prenatal diagnostic tests for BPD [16]. In a previous study, Ahmed et al. evaluated an ML technique also suitable for the analysis of other biological materials and created a helpful bedside point-of-care test approach for neonatal RDS [10]. According to the authors’ findings, following clinical validation, the use of ML-guided devices that can measure RDS biomarkers in real time may be used to direct therapies for preterm infants exhibiting respiratory symptoms. Moreover, Raimondi et al. concentrated on AI-assisted analysis of lung ultrasonography and its capacity to correlate with respiratory status in critically ill neonates with RDS [11]. The authors constructed a dataset of scans for texturing and a correlation between the oxygenation status, the ultrasound findings, and the mean grayscale intensity was established by an ML model. They enrolled a cohort of neonates of different origins and varying degrees of respiratory distress, and they demonstrated a significant correlation between blood gas indices and the grayscale ML analysis [11]; however, the relatively small sample size, the heterogeneous etiology of the respiratory distress, and the variable postnatal age suggested that further research on this topic with larger datasets is warranted.

Regarding BPD, Dai et al. investigated the combination of genetic and clinical factors, where exome sequencing was carried out for preterm neonates and integrated with clinical aspects [12]. The authors demonstrated that by using ML for the genomic analysis they could predict the development of BPD with an accuracy of 90% [12]. Also, the combination of gastric aspirate after birth and clinical information analysis could predict BPD development with a sensitivity of 88% [16]. Besides, Leigh et al., in a retrospective analysis of the perinatal and the respiratory factors in a sample of preterm neonates, created an ML algorithm that, after testing and training, could predict BPD-free survival well in terms of accuracy [14]. An AI approach has been proposed using DL and image segmentation, that can predict the severity of BPD by analyzing the segmentation of the lungs in chest X-rays taken on the 28th day of oxygen delivery [17]. The benefits of the aforementioned algorithm included non-invasiveness, speed, and independence from the experience of neonatologists, whereas demonstrated strong prediction performance.

Moreover, research on BPD with ML predictive models has shown that long-term invasive ventilation is one of the most significant risk factors for BPD and longer hospital stays. ML models using long-term invasive ventilation data could predict extubation failure with significant accuracy [106–108]. The risk stratification for BPD is a specific area of interest, aiming to identify infants who may benefit from preventive measures like corticosteroids or treatment for specific morbidities such as PDA. The BPD Outcome Estimator is a predictive tool approved by the US National Institute of Child Health and Human Development useful in directing steroid treatment and family counseling [13]. The estimator was initially limited to White, Black, or Hispanic neonates, however, Patel et al. recently created a web application based on an ML system for extremely preterm neonates of Asian descent [15]. Nonetheless, the study's conclusions were limited because the method was tested on a small dataset, requiring further comprehensive and prospective validation before being used in clinical practice.

Apnea of prematurity, another common morbidity in preterm neonates, is either obstructive (caused by airway obstruction), central (caused by cessation of respiratory drive), or mixed (a combination of both). Bedside monitors are programmed to sound an alarm when detect a decreased respiratory effort due to a decrease in thoracic motion [109]. A substantial number of false positive episodes have been observed in clinical tests indicating that this approach can identify central apneas with suboptimal accuracy [110]. Varisco et al. created an ML-based improved apnea detection model to automatically identify real apnea using data from the electrocardiographic monitoring of neonates [111]. The authors concluded that the AI algorithm resulted in better detection of apneas compared to traditional approaches with fewer false alarms, and they also showed that breathing patterns were altered more often in neonates with more frequent central apneas [111]. Although AI may drastically alter routine clinical practice, given that alarm fatigue is a growing problem in NICUs putting neonates in danger of missing alarms, the lack of external validation, along with the small sample size represents serious flaws in the suggested methodology. Table 3 presents examples of the application of AI in neonatal respiratory diseases.

3.4. Ophthalmology

ML models have been also applied in ROP, which is a severe complication of prematurity and a major cause of childhood blindness in high- and middle-income countries (Table 4). ROP affects mainly extremely preterm (less than 28 weeks), very preterm (28–32 weeks), or very low-birthweight (1500 g) neonates [23]. Telemedicine and AI are being considered as potential diagnostic tools for ROP, given the dearth of ophthalmologists who can treat neonates with ROP. Gaussian mixture models are among several ML techniques, to diagnose and categorize ROP from retinal fundus pictures [22,23]. In a previous study, the i-ROP system was shown to have a 95% accuracy in classifying pre-plus and plus illness. This performance was significantly better than the performance achieved by nonexperts (81%) and comparable to that achieved by experts (92% to 96%) [22].

Table 3. Examples of the current evidence of artificial intelligence application in neonatal respiratory diseases.

Aim	References	Artificial Intelligence Method	Data-Set Analyzed	Outcome
RDS severity	Ahmed et al. [10]	Attenuated total reflectance Fourier transform infrared spectroscopy combined with ML, performing callibration of principal component and partial least squares regression model	Two RDS biomarkers, lecithin and sphingomyelin (L/S ratio)	A three-factor model of second derivative spectra best predicted L/S ratios across the full range (R^2 : 0.967; MSE: 0.014). The L/S ratios from 1.0 to 3.4 were predicted with a prediction interval of +0.29, -0.37 when using a second derivative spectra model and had a mean prediction interval of +0.26, -0.34 around the L/S 2.2 region Visual assessment correlated significantly with respiratory indexes with a strong interobserver agreement. The use of regions of interest in the grayscale analysis of lung ultrasonography scans revealed a strong connection with oxygenation indexes
	Raimondi et al. [11]	SVM regressor	Lung ultrasonography using grayscale analysis supported by both visual and computer aids	An algorithm combining birth weight, gestational age, and the sectoral analysis of the gastric aspirates resulted to a sensitivity of 88% and a specificity of 91% for early diagnosis of BPD
Prediction of BPD	Verder et al. [16]	SVM	Clinical and laboratory data	The predictive model for BPD, which combined the BPD risk score and basic clinical risk factors, showed better discrimination than the model that was only based on basic clinical features (AUROC, 0.915 vs. AUROC, 0.814, $p = 0.013$, respectively). The severe BPD predictive model had AUROC, 0.907 vs. AUROC, 0.826; $p = 0.016$
	Dai et al. [12]	Predictive models evaluated using AUROC	Clinical and genetic features	The performance of the model showed AUROC 0.921 and 0.899 for the training and the validation datasets, respectively
	Leigh et al. [14]	A final ensemble model using logistic regression and the AUROC	Perinatal factors and early postnatal respiratory support	During the XSEG-Net network's training, the dice and cross-entropy loss values were 0.9794 and 0.0146, respectively. The deep CNN model based on VGGNet had the promising prediction performance, with the accuracy, precision, sensitivity, and specificity reaching 95.58%, 95.61%, 95.67%, and 96.98%, respectively
	Xing et al. [17]	XSEG-Net model combining digital image processing and human-computer interaction	Chest X-ray images	Prediction improved with advancing postnatal age, increasing from a C statistic of 0.793 on Day 1 to a maximum of 0.854 on Day 28
	Laughon et al. [13]	Models using a C statistic and AUROC	Gestational age, birth weight, race, ethnicity, sex, respiratory support, and FiO ₂	Model had AUROC of 0.934, 0.850, and 0.757 for respiratory outcomes at post-menstrual age 36, 37, and 40 weeks, respectively. An interrelationship among racial/ethnic groups and the feasibility of extending the use of the Estimator to the Asian population was shown
	Patel et al. [15]	Random forest algorithm with AUROC	Three racial/ethnic options	The optimal ANN model used 13 parameters and achieved an AUROC of 0.87, comparing favorably with multivariate logistic regression. It compared well with the clinician's expertise
Extubation readiness	Mueller et al. [106]	A ML approach using ANNs, multivariate logistic regression and the AUROC	51 variables	The predictor correctly identified infants who would not survive extubation, according to the results
	Precup et al. [107]	ML method of SVM	Measures of cardiorespiratory variability	Algorithm performance had AUROC of 0.871, sensitivity 70.1%, and specificity 90%
Automated detection of apneas	Mikhno et al. [108]	ML approach	Clinical and laboratory factors	The apnea detection model returned the highest mean AUROC, both using leave-one-patient-out and 10-fold cross-validation (mean AUROC of 0.88 and 0.90, respectively)
	Varisco et al. [111]	Optimized algorithm for automated detection using logistic regression and the AUROC	47 characteristics were taken out of the oxygen saturation and ECG signals	

RDS, respiratory distress syndrome; ML, machine learning; L/S, lecithin/sphingomyelin; SVM, support vector machine; BPD, bronchopulmonary dysplasia; AUROC, area under the receiver operating characteristic curve; CNN, convolutional neural network; ANN, artificial neural networks; ECG, electrocardiography.

Furthermore, a DL automated score model was generated in a recent multicenter trial, to identify one of the features of the affected retina [28]. This study showed how a DL comprehensive screening platform may enhance screening accessibility and objective ROP diagnosis. In another large-scale multicenter trial, a different group of scientists created a DL method for predicting ROP and its severity [30]. Retinal images from the initial ROP screening and neonatal clinical risk variables were obtained to develop an AI predictive algorithm. When compared to the traditional ROP score, the DL-based system demonstrated comparable accuracy, while it was found more effective in identifying and interpreting abnormal signs than the classical ophthalmoscopy.

Table 4. Examples of the current evidence of artificial intelligence application in neonatal ophthalmology.

Aim	References	Artificial Intelligence Model	Data-Set Analyzed	Outcome
Automated diagnosis of ROP	Ataer-Cansizoglu et al. [22]	Computer-based image analysis system (i-ROP)	Retina image	When compared to the reference standard, the i-ROP system classified preplus and plus illness with 95% accuracy. This was comparable to the performance of the 3 individual experts (96%, 94%, 92%), and significantly higher than the mean performance of 31 nonexperts (81%)
	Redd et al. [28]	A DL system (i-ROP plus score) on a 1–9 scale	Retina image	The AUROC of 0.960 was found for the i-ROP severity score in identifying type 1 ROP. Establishing a threshold i-ROP score of 3 conferred 94% sensitivity, 79% specificity, 13% positive predictive value and 99.7% negative predictive value for type 1 ROP. The i-ROP DL vascular severity score and expert rank ordering of overall ROP severity revealed a strong correlation ($r = 0.93$; $p < 0.0001$)
	Wu et al. [30]	Two models, OC-Net and SE-Net of ROP. AUROC, accuracy, sensitivity, and specificity	Retina image	AUROC, accuracy, sensitivity, and specificity were 0.90, 52.8%, 100%, and 37.8%, respectively, for OC-Net and 0.87, 68.0%, 100%, and 46.6%, respectively, for SE-Net. In external validation, the AUROC, accuracy, sensitivity, and specificity were 0.94, 33.3%, 100%, and 7.5%, respectively, for OC-Net, and 0.88, 56.0%, 100%, and 35.3% for SE-Net, respectively
	Biten et al. [24]	Telemedicine diagnoses of all 3 image readers	Retina image	Ophthalmoscopy and telemedicine each had similar sensitivity for zone I disease (78% vs. 78%), plus disease (74% vs. 79%), and type 2 ROP (stage 3, zone I, or plus disease: 86% vs. 79%), but ophthalmoscopy was slightly more sensitive in identifying stage 3 disease (85% vs. 73%; $p = 0.004$)
	Brown et al. [25]	Deep CNN algorithm based on deep learning. Receiver operating characteristic analysis was performed	Retina image	The diagnosis of plus disease (as opposed to pre-plus disease or normal) had an average AUROC of 0.98, whereas the diagnosis of normal (as opposed to pre-plus disease or normal) was 0.94. The method achieved 93% sensitivity and 94% specificity for + illness detection. The sensitivity and specificity for identifying pre-plus illness or worse were 100% and 94%, respectively
	Taylor et al. [29]	An algorithm assessing plus illness and its usefulness for impartially tracking the advancement of ROP	Retina image	The median severity scores for each category were 1.1 (no ROP), 1.5 (mild ROP), 4.6 (type 2 and pre-plus), and 7.5 (treatment-requiring ROP) ($p < 0.001$)
	Campbell et al. [26]	AI-based quantitative severity scale for ROP and AUROC	Retina image	The AUROC for detection of treatment-requiring retinopathy of prematurity was 0.98, with 100% sensitivity and 78% specificity

ROP, retinopathy of prematurity; DL, deep learning; AUROC, area under the receiver operating characteristic curve; OC-Net, occurrence network; SE-Net, severity network; CNN, convolutional neural network.

Moreover, in previous studies, telemedicine has been compared with Binocular Indirect Ophthalmoscope, demonstrating that both techniques are equally sensitive in detecting zone disease, plus disease, and ROP, although Binocular Indirect Ophthalmoscope was more accurate in recognizing zone III and stage 3 ROP [24,27]. Besides, using DL algorithms, the accuracy of ROP examination was 94% for normal diagnosis and 98% for illness and diagnosis, outperforming ROP experts [25]. Finally, in previous studies, DL algorithms were constructed to estimate the clinical progression of the ROP by assigning vascular severity scores [29] or to detect disease requiring therapy with an accuracy of 98% [26]. Overall, introducing AI into ROP screening programs might improve access to care for secondary ROP prevention [26]; however, despite the encouraging results, more extensive external validation using additional multicenter datasets is necessary. Additionally, the development of more advanced ML algorithms may be able to provide more significant prognostic information regarding the accurate staging, zone, and disease.

3.5. Gastrointestinal System

Recently, an AI algorithm was created based on a large dataset about the clinical characteristics of neonates who developed intestinal perforation [112] (Table 5). The suggested algorithm evaluated various clinical data, including vital signs, radiologic findings, biomarkers, and laboratory results, and led to a more accurate and early prediction of intestinal perforation of preterm neonates compared to all other traditional ML methods [112]. Furthermore, regarding nutrition, a previous study in England demonstrated

that ML techniques can be used to evaluate nutritional practices that were found to be associated with body weight on discharge and the development of BPD [113]. Finally, Han et al. recently examined the potential application of AI to predict postnatal growth failure. Using a large dataset of very low birth weight neonates from several NICUs, ML models were created using a variety of methodologies, showing a strong predictive performance [114]. Nevertheless, the study’s findings were limited since it lacked crucial information about enteral and parenteral feeding.

Table 5. Examples of the current evidence of artificial intelligence application in neonatal gastrointestinal diseases, sepsis, and patent ductus arteriosus.

Aim	References	Artificial Intelligence Model	Data-Set Analyzed	Outcome
Gastrointestinal System				
Prediction of spontaneous intestinal perforation	Son et al. [112]	ANNs and AUROC	Clinical data	The ANN models showed AUROC of 0.8832 for predicting intestinal perforation associated with necrotizing enterocolitis and 0.8797 for spontaneous perforation
Prediction of postnatal growth failure	Han et al. [114]	ML models were built using four different techniques XGB, random forest, SVM, and CNN to compare against the multiple logistic regression model	Clinical data	When compared with multiple logistic regression, XGB showed a significantly higher AUROC ($p = 0.03$) for Day 7, which was the primary performance metric. Using optimal cut-off points, for Day 7, XGB showed better performances in terms of AUROC (0.74), accuracy (0.68)
Sepsis				
Prediction of EOS	Adam et al. [20]	ML in form of a random forest classifier	Risk factors, clinical signs and biomarkers	The full model achieved an area under the receiver operating characteristic curve (AUROC) of 83.41% and an area under the precision recall curve 28.42%. The predictive performance of the model with risk factors alone was comparable with random
Prediction of LOS	Cabrera-Quiros et al. [21]	Three popular ML techniques (naive Bayes, closest mean classifier, and logistic regression)	ECG and respiration data (heart rate variability, respiration, body motion)	Using a combination of all features, classification of LOS and C showed a mean accuracy of 0.79 ± 0.12 and mean precision rate of 0.82 ± 0.18 3 h before the onset of sepsis
Patent ductus arteriosus				
Detection of PDA	Na et al. [19]	Algorithms including random forest, decision tree-based theory, L-GBM, low-bias model, feedforward ANN, SVM, using multiple logistic regression	Database of risk factors	L-GBM achieved the highest accuracy at predicting PDA (0.77), AUROC (0.82) and specificity (0.84), and logistic regression performed best with sensitivity (0.85). The random forest model achieved the best accuracy (0.85), AUROC (0.82) and sensitivity (0.97) in determining PDA therapy
	Gomez-Quintana et al. [18]	Clinical decision support tool based on ML	Heart sounds	The developed system reached an AUROC of 77% at detecting PDA. The obtained results for PDA detection compare favourably with the level of accuracy achieved by an experienced neonatologist when assessed on the same cohort

ANN, artificial neural networks; AUROC, area under the receiver operating characteristic curve; ML, machine learning; XGB, extreme gradient boosting; SVM, support vector machine; CNN, convolutional neural network; EOS, early-onset sepsis; LOS, late-onset sepsis; PDA, patent ductus arteriosus; L-GBM, light gradient boosting machine; ANN, artificial neural networks.

3.6. Sepsis

Early and late-onset neonatal sepsis is a major cause of infant mortality and morbidity [115]. Diagnosing neonatal sepsis and starting antibiotics is challenging in clinical practice, which emphasizes the need for a comprehensive approach. Previous studies have explored the role of heart rate variability in predicting early-onset sepsis with an accuracy of 64–94% [20]. Also, regarding the detection of late-onset sepsis, ML decision algorithms have utilized clinical and laboratory biomarkers obtaining an optimal accuracy and a mean precision rate of 0.82 3 h before the onset of sepsis [21] (Table 5).

3.7. Patent Ductus Arteriosus

The ductus arteriosus which is patent during the intrauterine life may have significant hemodynamic consequences in preterm neonates and is associated with higher rates of morbidity and mortality. Therefore, it should be assessed whether closing the PDA could

increase survival chances relative to the risk of side effects [116]. ML techniques have been developed for the detection of PDA from electronic health records [19] and auscultation records [18] (Table 5). This resulted in an accuracy of 76% for the prediction of PDA in very low birth weight infants based on the analysis of 47 perinatal factors using 5 different ML techniques [19] and 74% for the analysis of 250 auscultation records [18].

3.8. Dermatology

Infantile hemangiomas (IH) may present at birth and usually grow quickly between the ages of one and three months, so it's critical to diagnose the condition at an early age to avoid complications [117]. In a recent work by Zhang et al., a CNN was used to identify IH using clinical photos, reporting a diagnostic accuracy rate of 91.7%, which was even higher when restricting the analysis to the facial region [118]. This study showed that AI algorithms may be used for non-standardized photos, indicating their relevance to the real-world clinical context [118]. Future research on IH diagnosis will need to develop algorithms that can distinguish between different diseases instead of using a binary classifier, in addition to the capacity to categorize IH risk.

Although there is limited research on AI's application for pediatric dermatology issues, studies have examined adult illnesses that frequently affect pediatric patients. Atopic dermatitis is a recurrent condition that usually starts early in life [119]. A CNN was recently created by Guimaraes et al. to examine multiphoton tomography data for atopic dermatitis, with a diagnostic accuracy of 97% [120]. Furthermore, for the diagnosis of atopic dermatitis, De Guzman et al. created a multi-model, multi-level approach that produced a higher average confidence level (68.37% vs. 63.01%, respectively) than a single-model method [121]. Gustafson et al. used a phenotypic method based on ML to identify patients with atopic dermatitis in 2017. The system achieved a high positive prediction value and sensitivity by combining code information with the electronic health record collection. These findings show how ML and natural language processing can be used for EHR-based phenotyping [122]. The majority of current research uses adult database photos, where patient age is not clearly distinguished. This may cause biases in algorithms that are used for purposes other than clearly stating the ages for whom they are intended. Besides, a method based on deep neural networks was used by Han et al. to classify extremely rare skin lesions and distinguish between eczema and other infectious skin disorders. The authors also demonstrated that distinguishing between inflammatory and infectious causes could help with treatment options [123]. Moreover, a support-vector-machine-based image processing technique was developed for hand eczema segmentation and reported better results compared to other sophisticated approaches that were also tested [124] (Table 6).

Table 6. Examples of the current evidence of artificial intelligence application in neonatal dermatology.

Aim	References	Artificial Intelligence Model	Data-Set Analyzed	Outcome
Infantile hemangiomas	Zhang et al. [118]	Artificial intelligence algorithm	Clinical images	The algorithm achieved a 91.7% overall accuracy in the diagnosis of facial infantile hemangiomas
Atopic dermatitis	Guimaraes et al. [120]	CNN	Images combining both morphological and metabolic information	The algorithm correctly diagnosed atopic dermatitis in $97.0 \pm 0.2\%$ of all images presenting living cells. For diagnosis sensitivity was 0.966 ± 0.003 and specificity 0.977 ± 0.003
	De Guzman et al. [121]	A multi-model, multi-level system using the ANN architecture		When evaluating eczema against non-eczema instances, the system's average confidence level was 68.37%, compared to 63.01% for the single level, or single model system

Table 6. Cont.

Aim	References	Artificial Intelligence Model	Data-Set Analyzed	Outcome
	Gustafson et al. [122]	ML-based phenotype algorithm, using the electronic health record, combined in a lasso logistic regression	Coded information extracted from encounter notes	The algorithm achieved high positive predictive value and sensitivity. These results demonstrate the utility of natural language processing and ML for electronic health record-based phenotyping
Eczema	Han et al. [123]	Artificial intelligence algorithm	Images of 174 disorders	The AUROC for malignancy detection were 0.928 ± 0.002 and 0.937 ± 0.004 . The AUROC of primary treatment suggestion were 0.828 ± 0.012 , 0.885 ± 0.006 , 0.885 ± 0.006 , and 0.918 ± 0.006 for steroids, antibiotics, antivirals, and antifungals, respectively. With the assistance of our algorithm, the sensitivity and specificity of clinicians for malignancy prediction were improved by 12.1% ($p < 0.0001$) and 1.1% ($p < 0.0001$), respectively
	Koller et al. [124]	An automatic image processing method for hand eczema segmentation based on SVM	Several experiments with different feature sets	The system achieved an F1 score of 58.6% for front sides of hands and 43.8% for back sides, which outperforms methods that were tested on the gold standard data set

CNN, convolutional neural network; ANN, artificial neural networks; AUROC, area under the receiver operating characteristic curve; SVM, support vector machine.

3.9. Miscellaneous

3.9.1. Vital Signs Monitoring

In previous studies, ML analysis has been developed to analyze physiologic data that are electronically captured as signal data to identify artifact patterns [125], predict neonatal morbidity [126], or identify late-onset sepsis [21]. An ML algorithm using electronically recorded vital signs within the first three hours of life, including heart rate and respiration rate of preterm neonates with a birth weight ≤ 2000 g and gestational age ≤ 34 weeks predicted overall morbidity with an accuracy of 91% [126]. Furthermore, Lyra et al. developed DL-based techniques that could result in a reliable, real-time assessment of crucial indicators, such as changes in body temperature [127]. Although the analysis proved difficult for several factors during the recording, the authors demonstrated the viability of using inexpensive, embedded graphics processing units to monitor neonates' temperatures in real-time, although more research is warranted to broaden the application of this technique in clinical settings [127] (Table 7).

Table 7. Examples of the current evidence of artificial intelligence application in neonatal miscellaneous domains.

Aim	References	Artificial Intelligence Model	Data-Set Analyzed	Outcome
Vital Signs Monitoring				
Detect artifacts	Tsien et al. [125]	Decision tree induction	Multiple physiologic data signals	Finding artefacts was possible by the integration of many signals by using a classification system on sets of values obtained from physiologic data streams PhysiScore provided higher accuracy prediction of overall morbidity (86% sensitive at 96% specificity) than other neonatal scoring systems. PhysiScore was particularly accurate at identifying infants with high morbidity related to specific complications (infection: 90% to 100%; cardiopulmonary: 96% to 100%)
Predict overall mortality	Saria et al. [126]	Prediction algorithm (PhysiScore) based on a physiological assessment score for preterm newborns	Apgar score and standard signals recorded noninvasively on admission	

Table 7. Cont.

Aim	References	Artificial Intelligence Model	Data-Set Analyzed	Outcome
Temperature detection	Lyra et al. [127]	A combination of DL-based algorithms and camera modalities	Thermographic recordings	The keypoint detector's validation revealed a mean average precision of 0.82. The evaluation of the temperature extraction revealed a mean absolute error of 0.55 °C
Neonatal Jaundice				
Detection of jaundice	Althnian et al. [128]	DL approach	Eye, skin, and fused images	Traditional models outperformed DL models with eyes and fused features, but DL model did better with skin photos The ML diagnostic ability to evaluate the risk for neonatal jaundice was 0.748 (AUROC). Important factors were maternal blood type, maternal age, gestational age, estimated birth weight, parity, full blood count, and maternal blood pressure
	Guedalia et al. [129]	ML using a combined data analysis approach with AUROC	Clinical data without serum bilirubin evaluation	

DL, deep learning; ML, machine learning; AUROC, area under the receiver operating characteristic curve.

3.9.2. Neonatal Jaundice

The application of ML and DL models was explored in a previous study investigating the potential of using a dataset made up of photos taken using a smartphone camera for the identification of neonatal jaundice in term and late preterm neonates. The authors used data from pictures of the skin and eyes to train a neural network to identify jaundice [128]. Furthermore, Guardalia et al. used an ML approach to analyze clinical data for a large neonatal population to develop a risk assessment tool for neonatal jaundice that did not rely on bilirubin readings, that performed well in the risk categorization of newborn jaundice [129] (Table 7).

3.10. Mortality

Even with the recent advances in neonatal care, preterm neonates are still very vulnerable to death because of their immature organ systems [130]. ML models have been developed for the prediction of neonatal mortality by exploring causative factors [32,38] (Table 8). A recent review including term and preterm neonates between the gestational ages of 22 and 40 weeks reported that neural networks, random forests, and logistic regression were common models developed by the investigators [131]. Among the included studies, only two studies finished external validation, five studies published calibration plots, five studies reported sensitivity and specificity of their models that ranged from 63 to 80% and 78 to 98% respectively, and eight reported accuracy that ranged from 58.3 to 97.0% [131]. Despite having 17 features, the best model overall was linear regression analysis [131]. Recent studies exploring the application of AI models in severely low birthweight and preterm neonatal populations reported an accuracy of 68.9–93.3% [34,35]. Among the several limitations of these studies was the lack of inclusion of vital parameters to depict dynamic changes, while gestational age, birth weight, and Apgar scores were the most significant variables in the models [36,37]. These limitations suggest that further implementation, calibration, and external validation of AI healthcare applications is warranted in future studies.

Table 8. Examples of the current evidence of artificial intelligence application in neonatal mortality.

Aim	References	Artificial Intelligence Model	Data-Set Analyzed	Outcome
Prediction of mortality	Podda et al. [38]	ML methods including ANN, using logistic regression models	Twelve easily collected perinatal variables	ANN had a slightly better discrimination than logistic regression. Using a cutoff of death probability of 0.5, logistic regression misclassified 1.2 percent more than ANN
	Ambalavanan et al. [32]	Logistic regression and neural network models	Twenty-eight routinely collected variables were selected and multiple scenarios were created	The prediction was best with scenario C (AUROC: 0.85 for regression; 0.84 for neural networks), compared with scenarios A and B
	Hsu et al. [35]	ML of RF, bagged classification, and regression tree model with AUROC compared with the conventional neonatal illness severity scoring systems	Clinical and laboratory data	RF model showed the highest AUROC (0.939) for the prediction of neonates with respiratory failure, and the bagged classification and regression tree model demonstrated the next best results (0.915). The AUCs of both models were significantly better than the traditional severity scoring systems
	Do et al. [34]	ML methods including ANN, RF, and SVM	Neonatal and maternal factors	The model performances of AUROC equaled Logistic regression 0.841, ANN 0.845, and RF 0.826. The exception was SVM 0.631
	Moreira et al. [36]	Model performance was assessed via AUROC	Accessible clinical variables, gathered in the first hour following delivery	The model consisted of three variables: birth weight, Apgar score at 5 min of age, and gestational age. This model had an AUROC of 76.9%, while birth weight and gestational age had an AUROC of 73.1% and 71.3%
	Nascimento et al. [37]	A linguistic fuzzy model with minimum of Mamdani inference method	Neonatal birth weight and gestational age at delivery	The results were compared with experts' opinions and the Fuzzy model was able to capture the expert knowledge with a strong correlation ($r = 0.96$)

ML, machine learning; ANN, artificial neural networks; AUROC, area under the receiver operating characteristic curve; RF, random forest; SVM, support vector machine.

4. Challenges, Limitations, and Future Perspectives of Artificial Intelligence in Neonatology

AI has been currently established as a useful component in several parts of neonatal care, to help physicians to provide improved, more effective, and safer care (Table 9). However, specific issues need to be addressed before the wide application of AI models. At first, healthcare providers need to improve their digital literacy, so that they can comprehend the fundamental principles and limitations of AI. That would help healthcare providers evaluate recently created AI tools and focus on their appropriate and safe application in clinical settings. Also, to develop and implement AI tools, cross-disciplinary, worldwide collaborations involving data scientists, computer scientists, healthcare providers, attorneys, and legislators are required. Additional drawbacks of AI include the lack of larger datasets to train the models, the heterogeneity of the data, generalizability problems, the lack of evidence-based guidelines for some diseases affecting neonates, and the cost. Applying AI to newborn care also involves addressing critical challenges such as the model's interpretability, the necessity of external validation to improve generalizability, and the necessity of appropriate evaluation of performance (Table 2).

Finally, there are serious ethical issues to be considered. Important decisions in neonatology are often accompanied by a complex and difficult ethical component, and multidisciplinary methods are necessary for advancement [132]. Informed consent, bias, safety, privacy of the patients, and allocation are among the ethical issues with AI applications in healthcare [133]. The use of AI in neonatology has become more challenging due to the necessary transparency, viability limitations, life-sustaining therapies, and various international restrictions [134]. To date, there hasn't been any reporting on how an ethics framework would be applied in neonatology yet.

Table 9. Challenges of artificial intelligence in neonatology.

Challenges of AI	Areas of Improvement
Quality of the dataset	AI tools require high-quality data to be trained. Studies should address limitation including small sample sizes, improper management of missing information, and heterogeneity evaluation in various demographic subsets
Model performance evaluation	Model performance should be continually evaluated on the entire dataset. Apart from the AUROC, additional performance metrics, such as the precision-recall curve, specificity/sensitivity, and calibration metrics should be assessed
Clinical impact and external validation	External validation is crucial because, as in different dataset or in clinical practice, the tool's performance may degrade due to an over-modeling of the training data. Also, the effectiveness of AI should be evaluated in terms of calibration and discrimination quality as well as patient outcomes and the clinical workflow
Comprehending Guidelines for critical evaluation, regulation, and oversight	Bed-side models should enhance intelligence, interpretability, and transparency methodological, critical appraisal, medicolegal problems, and necessary monitoring is required to guarantee the model's safe and effective usage
Ethics	Informed consent, bias, patient privacy, and allocation are among the ethical issues with health AI, and negotiating their solutions can be challenging. Important decisions in neonatology are often accompanied by a complex and difficult ethical component, and multidisciplinary methods are necessary for advancement

AI, artificial intelligence; AUROC, area under the receiver operating characteristic curve.

5. Conclusions

AI is becoming more and more important in healthcare services following our contemporary culture that moves toward automated decision support systems. The main advantage of using AI in healthcare is its ability to evaluate large volumes of medical data from multidisciplinary studies. This type of data is too complex for medical professionals to study quickly enough to find the diagnosis and determine a treatment plan. When trained with the right data, AI models function like human neurons and can quickly and accurately solve problems. Finding the appropriate treatment strategy requires accuracy and time, especially in intensive care units. When integrating AI models into NICU clinical practices including treatment and transport, trust is a crucial component. AI-based solutions can be used in NICUs mainly to confirm the current treatment plans rather than implement their recommendations. The current evidence regarding the application of AI in neonatology is encouraging, however, further research is warranted including retraining clinical trials and validating the outcomes to make AI algorithms more useful in the future.

Author Contributions: Conceptualization, D.R. and V.G.; methodology, D.R.; investigation, D.R.; resources, D.R.; data curation, D.R.; writing—original draft preparation, D.R.; writing—review and editing, M.B., K.K. and V.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Helm, J.M.; Swiergosz, A.M.; Haeberle, H.S.; Karnuta, J.M.; Schaffer, J.L.; Krebs, V.E.; Spitzer, A.I.; Ramkumar, P.N. Machine Learning and Artificial Intelligence: Definitions, Applications, and Future Directions. *Curr. Rev. Musculoskelet Med.* **2020**, *13*, 69–76. [[CrossRef](#)]
- Price, W.N., 2nd; Gerke, S.; Cohen, I.G. Potential Liability for Physicians Using Artificial Intelligence. *JAMA* **2019**, *322*, 1765–1766. [[CrossRef](#)]
- Sujith, A.V.L.N.; Sajja, G.S.; Mahalakshmi, V.; Nuhmani, S.; Prasanalakshmi, B. Systematic review of smart health monitoring using deep learning and Artificial intelligence. *Neurosci. Inform.* **2022**, *2*, 100028. [[CrossRef](#)]
- Esteva, A.; Robicquet, A.; Ramsundar, B.; Kuleshov, V.; DePristo, M.; Chou, K.; Cui, C.; Corrado, G.; Thrun, S.; Dean, J. A guide to deep learning in healthcare. *Nat. Med.* **2019**, *25*, 24–29. [[CrossRef](#)] [[PubMed](#)]
- Sarker, I.H. Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Comput. Sci.* **2021**, *2*, 420. [[CrossRef](#)] [[PubMed](#)]

6. Rubinger, L.; Gazendam, A.; Ekhtiari, S.; Bhandari, M. Machine learning and artificial intelligence in research and healthcare. *Injury* **2023**, *54*, S69–S73. [[CrossRef](#)]
7. Meskó, B.; Görög, M. A short guide for medical professionals in the era of artificial intelligence. *NPJ Digit. Med.* **2020**, *3*, 126. [[CrossRef](#)]
8. Liu, A.; Hahn, J.S.; Heldt, G.P.; Coen, R.W. Detection of neonatal seizures through computerized EEG analysis. *Electroencephalogr. Clin. Neurophysiol.* **1992**, *82*, 30–37. [[CrossRef](#)]
9. Gotman, J.; Flanagan, D.; Zhang, J.; Rosenblatt, B. Automatic seizure detection in the newborn: Methods and initial evaluation. *Electroencephalogr. Clin. Neurophysiol.* **1997**, *103*, 356–362. [[CrossRef](#)]
10. Ahmed, W.; Veluthandath, A.V.; Rowe, D.J.; Madsen, J.; Clark, H.W.; Postle, A.D.; Wilkinson, J.S.; Murugan, G.S. Prediction of Neonatal Respiratory Distress Biomarker Concentration by Application of Machine Learning to Mid-Infrared Spectra. *Sensors* **2022**, *22*, 1744. [[CrossRef](#)] [[PubMed](#)]
11. Raimondi, F.; Migliaro, F.; Verdoliva, L.; Gagnaniello, D.; Poggi, G.; Kosova, R.; Sansone, C.; Vallone, G.; Capasso, L. Visual assessment versus computer-assisted gray scale analysis in the ultrasound evaluation of neonatal respiratory status. *PLoS ONE* **2018**, *13*, e0202397. [[CrossRef](#)] [[PubMed](#)]
12. Dai, D.; Chen, H.; Dong, X.; Chen, J.; Mei, M.; Lu, Y.; Yang, L.; Wu, B.; Cao, Y.; Wang, J.; et al. Bronchopulmonary Dysplasia Predicted by Developing a Machine Learning Model of Genetic and Clinical Information. *Front. Genet.* **2021**, *12*, 689071. [[CrossRef](#)] [[PubMed](#)]
13. Laughon, M.M.; Langer, J.C.; Bose, C.L.; Smith, P.B.; Ambalavanan, N.; Kennedy, K.A.; Stoll, B.J.; Buchter, S.; Laptook, A.R.; Ehrenkranz, R.A.; et al. Prediction of bronchopulmonary dysplasia by postnatal age in extremely premature infants. *Am. J. Respir. Crit. Care Med.* **2011**, *183*, 1715–1722. [[CrossRef](#)] [[PubMed](#)]
14. Leigh, R.M.; Pham, A.; Rao, S.S.; Vora, F.M.; Hou, G.; Kent, C.; Rodriguez, A.; Narang, A.; Tan, J.B.C.; Chou, F.S. Machine learning for prediction of bronchopulmonary dysplasia-free survival among very preterm infants. *BMC Pediatr.* **2022**, *22*, 542. [[CrossRef](#)]
15. Patel, M.; Sandhu, J.; Chou, F.S. Developing a machine learning-based tool to extend the usability of the NICHD BPD Outcome Estimator to the Asian population. *PLoS ONE* **2022**, *17*, e0272709. [[CrossRef](#)] [[PubMed](#)]
16. Verder, H.; Heiring, C.; Ramanathan, R.; Scoutaris, N.; Verder, P.; Jessen, T.E.; Hoskuldsson, A.; Bender, L.; Dahl, M.; Eschen, C.; et al. Bronchopulmonary dysplasia predicted at birth by artificial intelligence. *Acta Paediatr.* **2021**, *110*, 503–509. [[CrossRef](#)]
17. Xing, W.; He, W.; Li, X.; Chen, J.; Cao, Y.; Zhou, W.; Shen, Q.; Zhang, X.; Ta, D. Early severity prediction of BPD for premature infants from chest X-ray images using deep learning: A study at the 28th day of oxygen inhalation. *Comput. Methods Programs Biomed.* **2022**, *221*, 106869. [[CrossRef](#)]
18. Gomez-Quintana, S.; Schwarz, C.E.; Shelevytsky, I.; Shelevytska, V.; Semenova, O.; Factor, A.; Popovici, E.; Temko, A. A Framework for AI-Assisted Detection of Patent Ductus Arteriosus from Neonatal Phonocardiogram. *Healthcare* **2021**, *9*, 169. [[CrossRef](#)]
19. Na, J.Y.; Kim, D.; Kwon, A.M.; Jeon, J.Y.; Kim, H.; Kim, C.R.; Lee, H.J.; Lee, J.; Park, H.K. Artificial intelligence model comparison for risk factor analysis of patent ductus arteriosus in nationwide very low birth weight infants cohort. *Sci. Rep.* **2021**, *11*, 22353. [[CrossRef](#)]
20. Adam, J.; Rupprecht, S.; Kunstler, E.C.S.; Hoyer, D. Heart rate variability as a marker and predictor of inflammation, nosocomial infection, and sepsis—A systematic review. *Auton. Neurosci.* **2023**, *249*, 103116. [[CrossRef](#)]
21. Cabrera-Quiros, L.; Kommers, D.; Wolvers, M.K.; Oosterwijk, L.; Arents, N.; van der Sluijs-Bens, J.; Cottaar, E.J.E.; Andriessen, P.; van Pul, C. Prediction of Late-Onset Sepsis in Preterm Infants Using Monitoring Signals and Machine Learning. *Crit. Care Explor.* **2021**, *3*, e0302. [[CrossRef](#)] [[PubMed](#)]
22. Ataer-Cansizoglu, E.; Bolon-Canedo, V.; Campbell, J.P.; Bozkurt, A.; Erdogmus, D.; Kalpathy-Cramer, J.; Patel, S.; Jonas, K.; Chan, R.V.; Ostmo, S.; et al. Computer-Based Image Analysis for Plus Disease Diagnosis in Retinopathy of Prematurity: Performance of the “i-ROP” System and Image Features Associated With Expert Diagnosis. *Transl. Vis. Sci. Technol.* **2015**, *4*, 5. [[CrossRef](#)] [[PubMed](#)]
23. Barrero-Castillero, A.; Corwin, B.K.; VanderVeen, D.K.; Wang, J.C. Workforce Shortage for Retinopathy of Prematurity Care and Emerging Role of Telehealth and Artificial Intelligence. *Pediatr. Clin. N. Am.* **2020**, *67*, 725–733. [[CrossRef](#)] [[PubMed](#)]
24. Biten, H.; Redd, T.K.; Moleta, C.; Campbell, J.P.; Ostmo, S.; Jonas, K.; Chan, R.V.P.; Chiang, M.F. Diagnostic Accuracy of Ophthalmoscopy vs. Telemedicine in Examinations for Retinopathy of Prematurity. *JAMA Ophthalmol.* **2018**, *136*, 498–504. [[CrossRef](#)]
25. Brown, J.M.; Campbell, J.P.; Beers, A.; Chang, K.; Ostmo, S.; Chan, R.V.P.; Dy, J.; Erdogmus, D.; Ioannidis, S.; Kalpathy-Cramer, J.; et al. Automated Diagnosis of Plus Disease in Retinopathy of Prematurity Using Deep Convolutional Neural Networks. *JAMA Ophthalmol.* **2018**, *136*, 803–810. [[CrossRef](#)] [[PubMed](#)]
26. Campbell, J.P.; Singh, P.; Redd, T.K.; Brown, J.M.; Shah, P.K.; Subramanian, P.; Rajan, R.; Valikodath, N.; Cole, E.; Ostmo, S.; et al. Applications of Artificial Intelligence for Retinopathy of Prematurity Screening. *Pediatrics* **2021**, *147*, e2020016618. [[CrossRef](#)] [[PubMed](#)]
27. Chiang, M.F.; Melia, M.; Buffenn, A.N.; Lambert, S.R.; Recchia, F.M.; Simpson, J.L.; Yang, M.B. Detection of clinically significant retinopathy of prematurity using wide-angle digital retinal photography: A report by the American Academy of Ophthalmology. *Ophthalmology* **2012**, *119*, 1272–1280. [[CrossRef](#)] [[PubMed](#)]

28. Redd, T.K.; Campbell, J.P.; Brown, J.M.; Kim, S.J.; Ostmo, S.; Chan, R.V.P.; Dy, J.; Erdogmus, D.; Ioannidis, S.; Kalpathy-Cramer, J.; et al. Evaluation of a deep learning image assessment system for detecting severe retinopathy of prematurity. *Br. J. Ophthalmol.* **2018**, *103*, 580–584. [[CrossRef](#)] [[PubMed](#)]
29. Taylor, S.; Brown, J.M.; Gupta, K.; Campbell, J.P.; Ostmo, S.; Chan, R.V.P.; Dy, J.; Erdogmus, D.; Ioannidis, S.; Kim, S.J.; et al. Monitoring Disease Progression With a Quantitative Severity Scale for Retinopathy of Prematurity Using Deep Learning. *JAMA Ophthalmol.* **2019**, *137*, 1022–1028. [[CrossRef](#)]
30. Wu, Q.; Hu, Y.; Mo, Z.; Wu, R.; Zhang, X.; Yang, Y.; Liu, B.; Xiao, Y.; Zeng, X.; Lin, Z.; et al. Development and Validation of a Deep Learning Model to Predict the Occurrence and Severity of Retinopathy of Prematurity. *JAMA Netw. Open* **2022**, *5*, e2217447. [[CrossRef](#)]
31. Ali, R.; Li, H.; Dillman, J.R.; Altaye, M.; Wang, H.; Parikh, N.A.; He, L. A self-training deep neural network for early prediction of cognitive deficits in very preterm infants using brain functional connectome data. *Pediatr. Radiol.* **2022**, *52*, 2227–2240. [[CrossRef](#)] [[PubMed](#)]
32. Ambalavanan, N.; Carlo, W.A.; Bobashev, G.; Mathias, E.; Liu, B.; Poole, K.; Fanaroff, A.A.; Stoll, B.J.; Ehrenkranz, R.; Wright, L.L.; et al. Prediction of death for extremely low birth weight neonates. *Pediatrics* **2005**, *116*, 1367–1373. [[CrossRef](#)]
33. Balta, D.; Kuo, H.; Wang, J.; Porco, I.G.; Morozova, O.; Schladen, M.M.; Cereatti, A.; Lum, P.S.; Della Croce, U. Characterization of Infants' General Movements Using a Commercial RGB-Depth Sensor and a Deep Neural Network Tracking Processing Tool: An Exploratory Study. *Sensors* **2022**, *22*, 7426. [[CrossRef](#)]
34. Do, H.J.; Moon, K.M.; Jin, H.S. Machine Learning Models for Predicting Mortality in 7472 Very Low Birth Weight Infants Using Data from a Nationwide Neonatal Network. *Diagnostics* **2022**, *12*, 625. [[CrossRef](#)] [[PubMed](#)]
35. Hsu, J.F.; Yang, C.; Lin, C.Y.; Chu, S.M.; Huang, H.R.; Chiang, M.C.; Wang, H.C.; Liao, W.C.; Fu, R.H.; Tsai, M.H. Machine Learning Algorithms to Predict Mortality of Neonates on Mechanical Intubation for Respiratory Failure. *Biomedicines* **2021**, *9*, 1377. [[CrossRef](#)] [[PubMed](#)]
36. Moreira, A.; Benvenuto, D.; Fox-Good, C.; Alayli, Y.; Evans, M.; Jonsson, B.; Hakansson, S.; Harper, N.; Kim, J.; Norman, M.; et al. Development and Validation of a Mortality Prediction Model in Extremely Low Gestational Age Neonates. *Neonatology* **2022**, *119*, 418–427. [[CrossRef](#)] [[PubMed](#)]
37. Nascimento, L.F.; Ortega, N.R. Fuzzy linguistic model for evaluating the risk of neonatal death. *Rev. Saude Publica* **2002**, *36*, 686–692. [[CrossRef](#)] [[PubMed](#)]
38. Podda, M.; Bacciu, D.; Micheli, A.; Bellu, R.; Placidi, G.; Gagliardi, L. A machine learning approach to estimating preterm infants survival: Development of the Preterm Infants Survival Assessment (PISA) predictor. *Sci. Rep.* **2018**, *8*, 13743. [[CrossRef](#)]
39. Schadl, K.; Vassar, R.; Cahill-Rowley, K.; Yeom, K.W.; Stevenson, D.K.; Rose, J. Prediction of cognitive and motor development in preterm children using exhaustive feature selection and cross-validation of near-term white matter microstructure. *Neuroimage Clin.* **2018**, *17*, 667–679. [[CrossRef](#)]
40. Smyser, C.D.; Dosenbach, N.U.; Smyser, T.A.; Snyder, A.Z.; Rogers, C.E.; Inder, T.E.; Schlaggar, B.L.; Neil, J.J. Prediction of brain maturity in infants using machine-learning algorithms. *Neuroimage* **2016**, *136*, 1–9. [[CrossRef](#)]
41. Sripada, K.; Bjuland, K.J.; Solsnes, A.E.; Haberg, A.K.; Grunewaldt, K.H.; Lohaugen, G.C.; Rimol, L.M.; Skranes, J. Trajectories of brain development in school-age children born preterm with very low birth weight. *Sci. Rep.* **2018**, *8*, 15553. [[CrossRef](#)]
42. Valavani, E.; Blesa, M.; Galdi, P.; Sullivan, G.; Dean, B.; Cruickshank, H.; Sitko-Rudnicka, M.; Bastin, M.E.; Chin, R.F.M.; MacIntyre, D.J.; et al. Language function following preterm birth: Prediction using machine learning. *Pediatr. Res.* **2022**, *92*, 480–489. [[CrossRef](#)] [[PubMed](#)]
43. Vassar, R.; Schadl, K.; Cahill-Rowley, K.; Yeom, K.; Stevenson, D.; Rose, J. Neonatal Brain Microstructure and Machine-Learning-Based Prediction of Early Language Development in Children Born Very Preterm. *Pediatr. Neurol.* **2020**, *108*, 86–92. [[CrossRef](#)] [[PubMed](#)]
44. Wee, C.Y.; Tuan, T.A.; Broekman, B.F.; Ong, M.Y.; Chong, Y.S.; Kwek, K.; Shek, L.P.; Saw, S.M.; Gluckman, P.D.; Fortier, M.V.; et al. Neonatal neural networks predict children behavioral profiles later in life. *Hum. Brain. Mapp.* **2017**, *38*, 1362–1373. [[CrossRef](#)]
45. Zimmer, V.A.; Glocker, B.; Hahner, N.; Eixarch, E.; Sanroma, G.; Gratacos, E.; Rueckert, D.; Gonzalez Ballester, M.A.; Piella, G. Learning and combining image neighborhoods using random forests for neonatal brain disease classification. *Med. Image Anal.* **2017**, *42*, 189–199. [[CrossRef](#)]
46. Rajpurkar, P.; Chen, E.; Banerjee, O.; Topol, E.J. AI in health and medicine. *Nat. Med.* **2022**, *28*, 31–38. [[CrossRef](#)] [[PubMed](#)]
47. Adegboro, C.O.; Choudhury, A.; Asan, O.; Kelly, M.M. Artificial Intelligence to Improve Health Outcomes in the NICU and PICU: A Systematic Review. *Hosp. Pediatr.* **2022**, *12*, 93–110. [[CrossRef](#)]
48. Choudhury, A.; Asan, O. Role of Artificial Intelligence in Patient Safety Outcomes: Systematic Literature Review. *JMIR Med. Inf.* **2020**, *8*, e18599. [[CrossRef](#)]
49. Choudhury, A.; Renjilian, E.; Asan, O. Use of machine learning in geriatric clinical care for chronic diseases: A systematic literature review. *JAMIA Open* **2020**, *3*, 459–471. [[CrossRef](#)]
50. Olive, M.K.; Owens, G.E. Current monitoring and innovative predictive modeling to improve care in the pediatric cardiac intensive care unit. *Transl. Pediatr.* **2018**, *7*, 120–128. [[CrossRef](#)]
51. Piccialli, F.; Somma, V.D.; Giampaolo, F.; Cuomo, S.; Fortino, G. A survey on deep learning in medicine: Why, how and when? *Inf. Fusion* **2021**, *66*, 111–137. [[CrossRef](#)]

52. Burt, J.R.; Torosdagli, N.; Khosravan, N.; RaviPrakash, H.; Mortazi, A.; Tissavirasingham, F.; Hussein, S.; Bagci, U. Deep learning beyond cats and dogs: Recent advances in diagnosing breast cancer with deep neural networks. *Br. J. Radiol.* **2018**, *91*, 20170545. [[CrossRef](#)] [[PubMed](#)]
53. Erickson, B.J.; Korfiatis, P.; Kline, T.L.; Akkus, Z.; Philbrick, K.; Weston, A.D. Deep Learning in Radiology: Does One Size Fit All? *J. Am. Coll. Radiol.* **2018**, *15*, 521–526. [[CrossRef](#)] [[PubMed](#)]
54. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaría, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *J. Big Data* **2021**, *8*, 53. [[CrossRef](#)] [[PubMed](#)]
55. Ghosh, A.; Sufian, A.; Sultana, F.; Chakrabarti, A.; De, D. Fundamental Concepts of Convolutional Neural Network. In *Recent Trends and Advances in Artificial Intelligence and Internet of Things*; Intelligent Systems Reference Library; Springer International Publishing: Cham, Switzerland, 2020; pp. 519–567.
56. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
57. Ker, J.; Wang, L.; Rao, J.; Lim, T. Deep Learning Applications in Medical Image Analysis. *IEEE Access* **2018**, *6*, 9375–9389. [[CrossRef](#)]
58. Kreimeyer, K.; Foster, M.; Pandey, A.; Arya, N.; Halford, G.; Jones, S.F.; Forshee, R.; Walderhaug, M.; Botsis, T. Natural language processing systems for capturing and standardizing unstructured clinical information: A systematic review. *J. Biomed. Inf.* **2017**, *73*, 14–29. [[CrossRef](#)]
59. Nadkarni, P.M.; Ohno-Machado, L.; Chapman, W.W. Natural language processing: An introduction. *J. Am. Med. Inf. Assoc.* **2011**, *18*, 544–551. [[CrossRef](#)] [[PubMed](#)]
60. Brinkmann, B.H.; Bower, M.R.; Stengel, K.A.; Worrell, G.A.; Stead, M. Large-scale electrophysiology: Acquisition, compression, encryption, and storage of big data. *J. Neurosci. Methods* **2009**, *180*, 185–192. [[CrossRef](#)] [[PubMed](#)]
61. Sheth, R.D.; Hobbs, G.R.; Mullett, M. Neonatal seizures: Incidence, onset, and etiology by gestational age. *J. Perinatol.* **1999**, *19*, 40–43. [[CrossRef](#)]
62. Williams, R.P.; Banwell, B.; Berg, R.A.; Dlugos, D.J.; Donnelly, M.; Ichord, R.; Kessler, S.K.; Lavelle, J.; Massey, S.L.; Hewlett, J.; et al. Impact of an ICU EEG monitoring pathway on timeliness of therapeutic intervention and electrographic seizure termination. *Epilepsia* **2016**, *57*, 786–795. [[CrossRef](#)]
63. Payne, E.T.; Zhao, X.Y.; Frndova, H.; McBain, K.; Sharma, R.; Hutchison, J.S.; Hahn, C.D. Seizure burden is independently associated with short term outcome in critically ill children. *Brain* **2014**, *137*, 1429–1438. [[CrossRef](#)]
64. Chapman, K.E.; Specchio, N.; Shinnar, S.; Holmes, G.L. Seizing control of epileptic activity can improve outcome. *Epilepsia* **2015**, *56*, 1482–1485. [[CrossRef](#)]
65. Murray, D.M.; Boylan, G.B.; Ali, I.; Ryan, C.A.; Murphy, B.P.; Connolly, S. Defining the gap between electrographic seizure burden, clinical expression and staff recognition of neonatal seizures. *Arch. Dis. Child Fetal. Neonatal. Ed.* **2008**, *93*, F187–F191. [[CrossRef](#)] [[PubMed](#)]
66. Shellhaas, R.A.; Clancy, R.R. Characterization of neonatal seizures by conventional EEG and single-channel EEG. *Clin. Neurophysiol.* **2007**, *118*, 2156–2161. [[CrossRef](#)] [[PubMed](#)]
67. Scher, M.S.; Alvin, J.; Gaus, L.; Minnigh, B.; Painter, M.J. Uncoupling of EEG-clinical neonatal seizures after antiepileptic drug use. *Pediatr. Neurol.* **2003**, *28*, 277–280. [[CrossRef](#)]
68. McCoy, B.; Hahn, C.D. Continuous EEG monitoring in the neonatal intensive care unit. *J. Clin. Neurophysiol.* **2013**, *30*, 106–114. [[CrossRef](#)] [[PubMed](#)]
69. Shellhaas, R.A. Continuous long-term electroencephalography: The gold standard for neonatal seizure diagnosis. *Semin. Fetal. Neonatal. Med.* **2015**, *20*, 149–153. [[CrossRef](#)]
70. Shellhaas, R.A.; Chang, T.; Tsuchida, T.; Scher, M.S.; Riviello, J.J.; Abend, N.S.; Nguyen, S.; Wusthoff, C.J.; Clancy, R.R. The American Clinical Neurophysiology Society’s Guideline on Continuous Electroencephalography Monitoring in Neonates. *J. Clin. Neurophysiol.* **2011**, *28*, 611–617. [[CrossRef](#)]
71. de Vries, L.S.; Toet, M.C. Amplitude integrated electroencephalography in the full-term newborn. *Clin. Perinatol.* **2006**, *33*, 619–632. [[CrossRef](#)]
72. de Vries, L.S.; Hellstrom-Westas, L. Role of cerebral function monitoring in the newborn. *Arch. Dis. Child. Fetal. Neonatal. Ed.* **2005**, *90*, F201–F207. [[CrossRef](#)] [[PubMed](#)]
73. Rakshasbhuvankar, A.; Rao, S.; Palumbo, L.; Ghosh, S.; Nagarajan, L. Amplitude Integrated Electroencephalography Compared With Conventional Video EEG for Neonatal Seizure Detection: A Diagnostic Accuracy Study. *J. Child. Neurol.* **2017**, *32*, 815–822. [[CrossRef](#)]
74. Appendino, J.P.; McNamara, P.J.; Keyzers, M.; Stephens, D.; Hahn, C.D. The impact of amplitude-integrated electroencephalography on NICU practice. *Can. J. Neurol. Sci.* **2012**, *39*, 355–360. [[CrossRef](#)] [[PubMed](#)]
75. Temko, A.; Lightbody, G. Detecting Neonatal Seizures With Computer Algorithms. *J. Clin. Neurophysiol.* **2016**, *33*, 394–402. [[CrossRef](#)]
76. O’Shea, A.; Lightbody, G.; Boylan, G.; Temko, A. Neonatal seizure detection from raw multi-channel EEG using a fully convolutional architecture. *Neural. Netw.* **2020**, *123*, 12–25. [[CrossRef](#)]
77. O’Shea, A.; Ahmed, R.; Lightbody, G.; Pavlidis, E.; Lloyd, R.; Pisani, F.; Marnane, W.; Mathieson, S.; Boylan, G.; Temko, A. Deep Learning for EEG Seizure Detection in Preterm Infants. *Int. J. Neural. Syst.* **2021**, *31*, 2150008. [[CrossRef](#)] [[PubMed](#)]

78. Pavel, A.M.; Rennie, J.M.; de Vries, L.S.; Blennow, M.; Foran, A.; Shah, D.K.; Pressler, R.M.; Kapellou, O.; Dempsey, E.M.; Mathieson, S.R.; et al. A machine-learning algorithm for neonatal seizure recognition: A multicentre, randomised, controlled trial. *Lancet Child Adolesc. Health* **2020**, *4*, 740–749. [[CrossRef](#)]
79. Stevenson, N.J.; Korotchikova, I.; Temko, A.; Lightbody, G.; Marnane, W.P.; Boylan, G.B. An automated system for grading EEG abnormality in term neonates with hypoxic-ischaemic encephalopathy. *Ann. Biomed. Eng.* **2013**, *41*, 775–785. [[CrossRef](#)] [[PubMed](#)]
80. Mathieson, S.R.; Stevenson, N.J.; Low, E.; Marnane, W.P.; Rennie, J.M.; Temko, A.; Lightbody, G.; Boylan, G.B. Validation of an automated seizure detection algorithm for term neonates. *Clin. Neurophysiol.* **2016**, *127*, 156–168. [[CrossRef](#)]
81. Ansari, A.H.; Pillay, K.; Dereymaeker, A.; Jansen, K.; Van Huffel, S.; Naulaers, G.; De Vos, M. A Deep Shared Multi-Scale Inception Network Enables Accurate Neonatal Quiet Sleep Detection With Limited EEG Channels. *IEEE J. Biomed. Health Inf.* **2022**, *26*, 1023–1033. [[CrossRef](#)]
82. Raurale, S.A.; Boylan, G.B.; Mathieson, S.R.; Marnane, W.P.; Lightbody, G.; O’Toole, J.M. Grading hypoxic-ischemic encephalopathy in neonatal EEG with convolutional neural networks and quadratic time-frequency distributions. *J. Neural. Eng.* **2021**, *18*, 046007. [[CrossRef](#)]
83. Moghadam, S.M.; Pinchefskey, E.; Tse, I.; Marchi, V.; Kohonen, J.; Kauppila, M.; Airaksinen, M.; Tapani, K.; Nevalainen, P.; Hahn, C.; et al. Building an Open Source Classifier for the Neonatal EEG Background: A Systematic Feature-Based Approach From Expert Scoring to Clinical Visualization. *Front. Hum. Neurosci.* **2021**, *15*, 675154. [[CrossRef](#)]
84. Matic, V.; Cherian, P.J.; Koolen, N.; Naulaers, G.; Swarte, R.M.; Govaert, P.; Van Huffel, S.; De Vos, M. Holistic approach for automated background EEG assessment in asphyxiated full-term infants. *J. Neural. Eng.* **2014**, *11*, 066007. [[CrossRef](#)] [[PubMed](#)]
85. Pavel, A.M.; O’Toole, J.M.; Proietti, J.; Livingstone, V.; Mitra, S.; Marnane, W.P.; Finder, M.; Dempsey, E.M.; Murray, D.M.; Boylan, G.B.; et al. Machine learning for the early prediction of infants with electrographic seizures in neonatal hypoxic-ischemic encephalopathy. *Epilepsia* **2023**, *64*, 456–468. [[CrossRef](#)] [[PubMed](#)]
86. Serag, A.; Blesa, M.; Moore, E.J.; Pataky, R.; Sparrow, S.A.; Wilkinson, A.G.; Macnaught, G.; Semple, S.I.; Boardman, J.P. Accurate Learning with Few Atlases (ALFA): An algorithm for MRI neonatal brain extraction and comparison with 11 publicly available methods. *Sci. Rep.* **2016**, *6*, 23470. [[CrossRef](#)]
87. Blesa, M.; Galdi, P.; Cox, S.R.; Sullivan, G.; Stoye, D.Q.; Lamb, G.J.; Quigley, A.J.; Thrippleton, M.J.; Escudero, J.; Bastin, M.E.; et al. Hierarchical Complexity of the Macro-Scale Neonatal Brain. *Cereb. Cortex.* **2021**, *31*, 2071–2084. [[CrossRef](#)]
88. De Vries, L.S.; Groenendaal, F.; van Haastert, I.C.; Eken, P.; Rademaker, K.J.; Meiners, L.C. Asymmetrical myelination of the posterior limb of the internal capsule in infants with periventricular haemorrhagic infarction: An early predictor of hemiplegia. *Neuropediatrics* **1999**, *30*, 314–319. [[CrossRef](#)]
89. Odding, E.; Roebroek, M.E.; Stam, H.J. The epidemiology of cerebral palsy: Incidence, impairments and risk factors. *Disabil. Rehabil.* **2006**, *28*, 183–191. [[CrossRef](#)] [[PubMed](#)]
90. Drougia, A.; Giapros, V.; Krallis, N.; Theocharis, P.; Nikaki, A.; Tzoufi, M.; Andronikou, S. Incidence and risk factors for cerebral palsy in infants with perinatal problems: A 15-year review. *Early Hum. Dev.* **2007**, *83*, 541–547. [[CrossRef](#)]
91. Gruber, N.; Galijasevic, M.; Regodic, M.; Grams, A.E.; Siedentopf, C.; Steiger, R.; Hammerl, M.; Haltmeier, M.; Gizewski, E.R.; Janjic, T. A deep learning pipeline for the automated segmentation of posterior limb of internal capsule in preterm neonates. *Artif. Intell. Med.* **2022**, *132*, 102384. [[CrossRef](#)]
92. Dean, B.; Ginnell, L.; Boardman, J.P.; Fletcher-Watson, S. Social cognition following preterm birth: A systematic review. *Neurosci. Biobehav. Rev.* **2021**, *124*, 151–167. [[CrossRef](#)] [[PubMed](#)]
93. Batalle, D.; Edwards, A.D.; O’Muircheartaigh, J. Annual Research Review: Not just a small adult brain: Understanding later neurodevelopment through imaging the neonatal brain. *J. Child Psychol. Psychiatry* **2018**, *59*, 350–371. [[CrossRef](#)] [[PubMed](#)]
94. Ball, G.; Aljabar, P.; Nongena, P.; Kennea, N.; Gonzalez-Cinca, N.; Falconer, S.; Chew, A.T.M.; Harper, N.; Wurie, J.; Rutherford, M.A.; et al. Multimodal image analysis of clinical influences on preterm brain development. *Ann. Neurol.* **2017**, *82*, 233–246. [[CrossRef](#)]
95. Galdi, P.; Blesa, M.; Stoye, D.Q.; Sullivan, G.; Lamb, G.J.; Quigley, A.J.; Thrippleton, M.J.; Bastin, M.E.; Boardman, J.P. Neonatal morphometric similarity mapping for predicting brain age and characterizing neuroanatomic variation associated with preterm birth. *Neuroimage Clin.* **2020**, *25*, 102195. [[CrossRef](#)] [[PubMed](#)]
96. Makropoulos, A.; Gousias, I.S.; Ledig, C.; Aljabar, P.; Serag, A.; Hajnal, J.V.; Edwards, A.D.; Counsell, S.J.; Rueckert, D. Automatic whole brain MRI segmentation of the developing neonatal brain. *IEEE Trans. Med. Imaging* **2014**, *33*, 1818–1831. [[CrossRef](#)] [[PubMed](#)]
97. Ding, Y.; Acosta, R.; Enguix, V.; Suffren, S.; Ortmann, J.; Luck, D.; Dolz, J.; Lodygensky, G.A. Using Deep Convolutional Neural Networks for Neonatal Brain Image Segmentation. *Front. Neurosci.* **2020**, *14*, 207. [[CrossRef](#)] [[PubMed](#)]
98. Shang, J.; Fisher, P.; Bauml, J.G.; Daamen, M.; Baumann, N.; Zimmer, C.; Bartmann, P.; Boecker, H.; Wolke, D.; Sorg, C.; et al. A machine learning investigation of volumetric and functional MRI abnormalities in adults born preterm. *Hum. Brain Mapp.* **2019**, *40*, 4239–4252. [[CrossRef](#)] [[PubMed](#)]
99. Chiarelli, A.M.; Sestieri, C.; Navarra, R.; Wise, R.G.; Caulo, M. Distinct effects of prematurity on MRI metrics of brain functional connectivity, activity, and structure: Univariate and multivariate analyses. *Hum. Brain Mapp.* **2021**, *42*, 3593–3607. [[CrossRef](#)] [[PubMed](#)]
100. Ball, G.; Aljabar, P.; Arichi, T.; Tusor, N.; Cox, D.; Merchant, N.; Nongena, P.; Hajnal, J.V.; Edwards, A.D.; Counsell, S.J. Machine-learning to characterise neonatal functional connectivity in the preterm brain. *Neuroimage* **2016**, *124*, 267–275. [[CrossRef](#)]

101. Song, Z.; Awate, S.P.; Licht, D.J.; Gee, J.C. Clinical neonatal brain MRI segmentation using adaptive nonparametric data models and intensity-based Markov priors. *Med. Image Comput. Comput. Assist. Interv.* **2007**, *10*, 883–890. [[CrossRef](#)]
102. Keunen, K.; Counsell, S.J.; Benders, M. The emergence of functional architecture during early brain development. *Neuroimage* **2017**, *160*, 2–14. [[CrossRef](#)]
103. Gao, W.; Lin, W.; Grewen, K.; Gilmore, J.H. Functional Connectivity of the Infant Human Brain: Plastic and Modifiable. *Neuroscientist* **2017**, *23*, 169–184. [[CrossRef](#)]
104. Li, Y.; Zhang, X.; Nie, J.; Zhang, G.; Fang, R.; Xu, X.; Wu, Z.; Hu, D.; Wang, L.; Zhang, H.; et al. Brain Connectivity Based Graph Convolutional Networks and Its Application to Infant Age Prediction. *IEEE Trans. Med. Imaging* **2022**, *41*, 2764–2776. [[CrossRef](#)] [[PubMed](#)]
105. Krishnan, M.L.; Wang, Z.; Aljabar, P.; Ball, G.; Mirza, G.; Saxena, A.; Counsell, S.J.; Hajnal, J.V.; Montana, G.; Edwards, A.D. Machine learning shows association between genetic variability in PPARG and cerebral connectivity in preterm infants. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 13744–13749. [[CrossRef](#)] [[PubMed](#)]
106. Mueller, M.; Wagner, C.L.; Annibale, D.J.; Hulsey, T.C.; Knapp, R.G.; Almeida, J.S. Predicting extubation outcome in preterm newborns: A comparison of neural networks with clinical expertise and statistical modeling. *Pediatr. Res.* **2004**, *56*, 11–18. [[CrossRef](#)] [[PubMed](#)]
107. Precup, D.; Robles-Rubio, C.A.; Brown, K.A.; Kanbar, L.; Kaczmarek, J.; Chawla, S.; Sant’Anna, G.M.; Kearney, R.E. Prediction of extubation readiness in extreme preterm infants based on measures of cardiorespiratory variability. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2012**, *2012*, 5630–5633. [[CrossRef](#)]
108. Mikhno, A.; Ennett, C.M. Prediction of extubation failure for neonates with respiratory distress syndrome using the MIMIC-II clinical database. In Proceedings of the 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, USA, 28 August–1 September 2012. [[CrossRef](#)]
109. Eichenwald, E.C.; Committee on Fetus and Newborn; Watterberg, K.L.; Aucott, S.; Benitz, W.E.; Cummings, J.J.; Goldsmith, J.; Poindexter, B.B.; Puopolo, K.; Stewart, D.L.; et al. Apnea of Prematurity. *Pediatrics* **2016**, *137*, e20153757. [[CrossRef](#)]
110. Amin, S.B.; Burnell, E. Monitoring apnea of prematurity: Validity of nursing documentation and bedside cardiorespiratory monitor. *Am. J. Perinatol.* **2013**, *30*, 643–648. [[CrossRef](#)] [[PubMed](#)]
111. Varisco, G.; Peng, Z.; Kommers, D.; Zhan, Z.; Cottaar, W.; Andriessen, P.; Long, X.; van Pul, C. Central apnea detection in premature infants using machine learning. *Comput Methods Programs Biomed.* **2022**, *226*, 107155. [[CrossRef](#)]
112. Son, J.; Kim, D.; Na, J.Y.; Jung, D.; Ahn, J.H.; Kim, T.H.; Park, H.K. Development of artificial neural networks for early prediction of intestinal perforation in preterm infants. *Sci. Rep.* **2022**, *12*, 12112. [[CrossRef](#)]
113. Greenbury, S.F.; Ougham, K.; Wu, J.; Battersby, C.; Gale, C.; Modi, N.; Angelini, E.D. Identification of variation in nutritional practice in neonatal units in England and association with clinical outcomes using agnostic machine learning. *Sci. Rep.* **2021**, *11*, 7178. [[CrossRef](#)]
114. Han, J.H.; Yoon, S.J.; Lee, H.S.; Park, G.; Lim, J.; Shin, J.E.; Eun, H.S.; Park, M.S.; Lee, S.M. Application of Machine Learning Approaches to Predict Postnatal Growth Failure in Very Low Birth Weight Infants. *Yonsei Med. J.* **2022**, *63*, 640–647. [[CrossRef](#)]
115. Shane, A.L.; Sanchez, P.J.; Stoll, B.J. Neonatal sepsis. *Lancet* **2017**, *390*, 1770–1780. [[CrossRef](#)] [[PubMed](#)]
116. El-Khuffash, A.; Bussmann, N.; Breatnach, C.R.; Smith, A.; Tully, E.; Griffin, J.; McCallion, N.; Corcoran, J.D.; Fernandez, E.; Looi, C.; et al. A Pilot Randomized Controlled Trial of Early Targeted Patent Ductus Arteriosus Treatment Using a Risk Based Severity Score (The PDA RCT). *J. Pediatr.* **2021**, *229*, 127–133. [[CrossRef](#)] [[PubMed](#)]
117. Krowchuk, D.P.; Frieden, I.J.; Mancini, A.J.; Darrow, D.H.; Blei, F.; Greene, A.K.; Annam, A.; Baker, C.N.; Frommelt, P.C.; Hodak, A.; et al. Clinical Practice Guideline for the Management of Infantile Hemangiomas. *Pediatrics* **2019**, *143*, e20183475. [[CrossRef](#)] [[PubMed](#)]
118. Zhang, A.J.; Lindberg, N.; Chamlin, S.L.; Haggstrom, A.N.; Mancini, A.J.; Siegel, D.H.; Drolet, B.A. Development of an artificial intelligence algorithm for the diagnosis of infantile hemangiomas. *Pediatr. Dermatol.* **2022**, *39*, 934–936. [[CrossRef](#)] [[PubMed](#)]
119. Drucker, A.M.; Wang, A.R.; Li, W.-Q.; Severson, E.; Block, J.K.; Qureshi, A.A. The Burden of Atopic Dermatitis: Summary of a Report for the National Eczema Association. *J. Investig. Dermatol.* **2017**, *137*, 26–30. [[CrossRef](#)] [[PubMed](#)]
120. Guimarães, P.; Batista, A.; Zieger, M.; Kaatz, M.; Koenig, K. Artificial Intelligence in Multiphoton Tomography: Atopic Dermatitis Diagnosis. *Sci. Rep.* **2020**, *10*, 7968. [[CrossRef](#)] [[PubMed](#)]
121. De Guzman, L.C.; Maglaque, R.P.C.; Torres, V.M.B.; Zapido, S.P.A.; Cordel, M.O. Design and Evaluation of a Multi-model, Multi-level Artificial Neural Network for Eczema Skin Lesion Detection. In Proceedings of the 2015 3rd International Conference on Artificial Intelligence, Modelling and Simulation (AIMS), Kota Kinabalu, Malaysia, 2–4 December 2015; pp. 42–47.
122. Gustafson, E.; Pacheco, J.; Wehbe, F.; Silverberg, J.; Thompson, W. A Machine Learning Algorithm for Identifying Atopic Dermatitis in Adults from Electronic Health Records. In Proceedings of the 2017 IEEE International Conference on Healthcare Informatics (ICHI), Park City, UT, USA, 23–26 August 2017; pp. 83–90.
123. Han, S.S.; Park, I.; Eun Chang, S.; Lim, W.; Kim, M.S.; Park, G.H.; Chae, J.B.; Huh, C.H.; Na, J.-I. Augmented Intelligence Dermatology: Deep Neural Networks Empower Medical Professionals in Diagnosing Skin Cancer and Predicting Treatment Options for 134 Skin Disorders. *J. Investig. Dermatol.* **2020**, *140*, 1753–1761. [[CrossRef](#)]
124. Koller, T.; Navarini, A.; von der Brück, T.; Pouly, M.; Schnürle, S. On using Support Vector Machines for the Detection and Quantification of Hand Eczema. In Proceedings of the 9th International Conference on Agents and Artificial Intelligence, Porto, Portugal, 24–26 February 2017; pp. 75–84.

125. Tsien, C.L.; Kohane, I.S.; McIntosh, N. Multiple signal integration by decision tree induction to detect artifacts in the neonatal intensive care unit. *Artif. Intell. Med.* **2000**, *19*, 189–202. [[CrossRef](#)]
126. Saria, S.; Rajani, A.K.; Gould, J.; Koller, D.; Penn, A.A. Integration of early physiological responses predicts later illness severity in preterm infants. *Sci. Transl. Med.* **2010**, *2*, 48ra65. [[CrossRef](#)]
127. Lyra, S.; Rixen, J.; Heimann, K.; Karthik, S.; Joseph, J.; Jayaraman, K.; Orlikowsky, T.; Sivaprakasam, M.; Leonhardt, S.; Hoog Antink, C. Camera fusion for real-time temperature monitoring of neonates using deep learning. *Med. Biol. Eng. Comput* **2022**, *60*, 1787–1800. [[CrossRef](#)]
128. Althnian, A.; Almanea, N.; Aloboud, N. Neonatal Jaundice Diagnosis Using a Smartphone Camera Based on Eye, Skin, and Fused Features with Transfer Learning. *Sensors* **2021**, *21*, 7038. [[CrossRef](#)]
129. Guedalia, J.; Farkash, R.; Wasserteil, N.; Kasirer, Y.; Rottenstreich, M.; Unger, R.; Grisaru Granovsky, S. Primary risk stratification for neonatal jaundice among term neonates using machine learning algorithm. *Early Hum. Dev.* **2022**, *165*, 105538. [[CrossRef](#)] [[PubMed](#)]
130. Pearlman, S.A. Advancements in neonatology through quality improvement. *J. Perinatol.* **2022**, *42*, 1277–1282. [[CrossRef](#)] [[PubMed](#)]
131. Mangold, C.; Zoretic, S.; Thallapureddy, K.; Moreira, A.; Chorath, K.; Moreira, A. Machine Learning Models for Predicting Neonatal Mortality: A Systematic Review. *Neonatology* **2021**, *118*, 394–405. [[CrossRef](#)]
132. Mercurio, M.R.; Cummings, C.L. Critical decision-making in neonatology and pediatrics: The I-P-O framework. *J. Perinatol.* **2021**, *41*, 173–178. [[CrossRef](#)] [[PubMed](#)]
133. Katznelson, G.; Gerke, S. The need for health AI ethics in medical school education. *Adv. Health. Sci. Educ. Theory Pract.* **2021**, *26*, 1447–1458. [[CrossRef](#)]
134. Lin, M.; Vitcov, G.G.; Cummings, C.L. Moral equivalence theory in neonatology. *Semin. Perinatol.* **2022**, *46*, 151525. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.